



TRUSTEE

TRUSTEE D2.1 Live doc conceptualisation, use cases and system architecture V1

Work package	WP2: Conceptualisation, Use cases and System Architecture
Task	Task 2.1: SoA analysis, technological requirements, and benchmarking of best practices Task 2.2: Legal framework for data manipulation and AI processing Task 2.3: Reference Use Cases definition and requirements Task 2.4: Reference Architecture and system specifications Task 2.5: Federated Data Model for Data Leakage
Authors	Glossary: Andrej Grguric – ENT, Ivana Stupar – ENT, Regueiro Senderos, Cristina – TECNALIA, Dimitra Papatsaroucha – HMU, Sofiane Lagraa – FUJITSU, Nikolaus Forgó – UNIVIE, Katarzyna Barud – UNIVIE, Hande Özkayagan-Prändl – UNIVIE, Clara Saillant – UNIVIE, Alexandru Nicolin – ISS, Marco Di Girolamo – HPE, Evangelos Markakis – HMU, Maurizio Martignano – UCSC, Giulia Rutto – RINA Part I: Vasilis Pikoulis – ATHENA, Christos Mavrokefalidis – ATHENA, Christos Anagnostopoulos – ATHENA, Alexandros Gkillas – ATHENA, Nikos Piperigkos – ATHENA, Aris Lalos – ATHENA, Ilias Politis – IQB, Vassiliki Sambazioti – IQB, Athanasios Grammatopoulos – IQB, Maria Liatsou – IQB, Konstantinos Kefalas – IQB, Alexandros Fotiadis – IQB, Sofiane Lagraa – FUJITSU, Marisa Escalante Martinez – TECNALIA, Cristina Regueiro Senderos – TECNALIA, Iñaki Seco – TECNALIA, Santiago de Diego – TECNALIA, Maria Zotou – ACCELI, Pantelis Velanas - ACCELI, Giulia Rutto – RINA, Andrej Grguric – ENT, Ivana Stupar – ENT, Nikolaus Forgó – UNIVIE, Katarzyna Barud – UNIVIE, Hande Özkayagan-Prändl – UNIVIE, Clara Saillant – UNIVIE, Niskanen Ilkka – VTT, Alberto Terzi – HPE, Nicola Di Monte – HPE, Pasquale Di Dio – HPE, Sofiane Lagraa – FUJITSU, Moussa Ouedraogo – FUJITSU, Sven Rasmusen – FUJITSU, Marios Siganos – K3Y, Savvas Ouzounidis – K3Y, George Nakas – K3Y, Panagiotis Radoglou-Grammatikis – K3Y Part II: Dimitra Papatsaroucha – HMU, Evangelos Markakis – HMU, Konstantina Pityanou – HMU, Katarzyna Barud – UNIVIE, Maurizio Martignano – UCSC, Velislava Hillman – EPL, Niskanen Ilkka – VTT, Alexandru Nicolin – ISS, Emmanouil G. Spanakis – FORTH, Peter Rosengren – CHECKWATT, Hande Özkayagan-Prändl – UNIVIE, Clara Saillant – UNIVIE Part III: Dimitra Papatsaroucha – HMU, Kalliopi Meladaki – HMU, Velislava Hillman – EPL, Katarzyna Barud – UNIVIE, Nikolaus Forgó – UNIVIE, Katarzyna Barud – UNIVIE, Hande Özkayagan-Prändl – UNIVIE, Clara Saillant – UNIVIE Part IV: Ilias Politis – IQB, Vassiliki Sambazioti – IQB, Athanasios Grammatopoulos – IQB, Maria Liatsou – IQB, Konstantinos Kefalas – IQB, Alexandros Fotiadis – IQB, Dimitra Papatsaroucha – HMU, Andrej Grguric – ENT, Vasilis Pikoulis – ATHENA, Aris Lalos – ATHENA, Sofiane Lagraa – FUJITSU, Marisa Escalante Martinez – TECNALIA, Cristina Regueiro Senderos – TECNALIA, Maria Zotou – ACCELI, Pantelis Velanas - ACCELI, Stergios Zissakis - ACCELI, Stelios Vourakis - ACCELI, Giulia Rutto – RINA, Athanasios Grammatopoulos – IQB, Katarzyna Barud – UNIVIE, Maurizio Martignano – UCSC, Niskanen Ilkka – VTT, Emmanouil G. Spanakis – FORTH, Cristina Farmaki - FORTH Part V: Andrej Grguric – ENT, Ivana Stupar – ENT, Dimitra Papatsaroucha – HMU, Evangelos Markakis – HMU, Xenofon Vurakis – HMU, Vasilis Pikoulis – ATHENA, Aris Lalos –

	ATHENA, Sofiane Lagraa – FUJITSU, Marisa Escalante Martinez – TECNALIA, Cristina Regueiro Senderos – TECNALIA, Maria Zotou – ACCELI, Pantelis Velanas - ACCELI, Sterios Zissakis - ACCELI, Stelios Vourakis - ACCELI, Anastatios Nomikos - ACCELI, Giulia Rutto – RINA, Nikos Papachatzakis – ADR, Niskanen Ilkka – VTT, Emmanouil G. Spanakis – FORTH, Cristina Farmaki - FORTH
Dissemination level	Public (PU)
Status	Final
Due date	30/06/2023
Document date	30/06/23
Version number	1.0
 Funded by the European Union	<i>TRUSTEE project has received funding from the European Union's Horizon EUROPE research & innovation programme under grant agreement No 101070214.</i> <i>Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.</i>

Table of Contents	3
Index of Figures.....	10
Index of Tables	16
Executive Summary	41
Introduction	41
Deliverable Structure	41
TRUSTEE Manifesto	43
Part I: State of the Art.....	45
Part I Summary	45
Part I Introduction.....	45
Secure and Trusted Platforms for Data Exchange	47
Threat Landscape	47
Procedures, Policies, Technical Guidelines, Tools and Recommendations for Data Privacy and Confidentiality	53
Trustworthiness-by-Design Frameworks for Monitoring, Auditing and Conformity Check	56
Security and Trust Managers	60
Semantic Ontologies (ACC)	63
Legal Foundation on Data Privacy and Security in TRUSTEE.....	72
Fundamental Rights and Primary Law	72
Data Protection Legal Framework – Secondary Legislation	74
Enabling Exchange of Non-personal Data	100
Legal Framework for Data Governance and Access Included in Data Governance Act, Data Act (Proposal) and Open Data Directive	101
Regulating AI – Analysis of Initial Framework for AI System Providers and Users	110
Cybersecurity Legal Framework and Requirements for Information Systems Providers	124
eIDAS Regulation and Requirements	128
Digital Services Act and Requirements for Intermediary Services	132
P2B Regulation	134
B2B and B2C Data Sharing Principles and Contractual Arrangements	136

Blockchain Services – Relevant Considerations for Blockchain Providers and Users	139
Intellectual Property Law Framework	143
Homomorphic-capable Self-sovereign Framework	163
Homomorphic Encryption for Data Privacy and Usability	163
Identification, Authentication, and Authorization Based on SSI Framework	178
HE Combined with SSI	184
eIDAS over SSI Utilizing the IOTA Framework	185
Open and FAIR Framework for Extracted Knowledge	188
Concept	188
Knowledge Repository and External Tools	188
IPFS Overview	189
Review of the Use-Cases	191
Privacy, Anti-Censorship and Immutability	192
Limitations	192
Private IPFS	193
xAI Techniques Focusing on Transparency, Trustworthiness and Federated ML	193
Explainable AI Overview	193
XAI Applications in the Energy	195
XAI Applications in Automotive	196
XAI Applications in HEALTH	196
XAI Applications in EDUCATION	197
XAI Applications in SPACE	198
XAI BY DESIGN	199
Lessons Learnt	200
Privacy and Security in Federated Learning: Challenges, Solutions and Applications	201
Existing FL Frameworks	202
Federation Mechanisms for I.I.D. and Non-I.I.D. Data Models	210
Data Leakage in Federated Learning	221
Mitigation Mechanisms with Emphasis in HE-enabled FL	225
FL Applications Within TRUSTEE	227

SOTA on TRUSTEE-related Reference Architectures	250
Reference Architecture Model, Reference Architecture, Concrete Architecture, and System Definitions	250
Relevant Standards	252
Reference Architecture Models, with Industrial and Cross-Sector Focus	254
Reference Architecture Models, with a Focus on Data Spaces	261
Reference Architectures, with IoT Focus	263
Reference Architectures, with Digital Public Services Focus	264
Reference Architectures, with Big Data Focus	268
Reference Architectures, with a Focus on Data Space	269
EU AI Framework	275
TRUSTEE Landscape	277
Related EU Projects	277
Initiatives	288
Part II: Use Cases	289
Part II Executive Summary	289
Part II Introduction	289
Review of Requirements with Pilots – 1 st Iteration	289
Energy	289
Data Summary	289
Making Data Accessible	290
Data Security and Protection	290
Ethical Requirements	290
Other Issues	291
Health	291
Data Summary	291
Making Data Accessible	291
Data Security and Protection	293
Ethical Requirements	294
Other Issues	294
Education	294

Data Summary.....	294
Making Data Accessible.....	296
Data Security and Protection	296
Ethical Requirements	297
Other Issues	297
Automotive	297
Data Summary.....	297
Making Data Accessible.....	297
Data Security and Protection	298
Ethical Requirements	298
Other Issues	298
Space.....	298
Data Summary.....	298
Making Data Accessible.....	299
Data Security and Protection	299
Ethical Requirements	299
Other Issues	299
Multi-disciplinary Data Exchange.....	299
Data Summary.....	299
Making Data Accessible.....	300
Data Security and Protection	300
Ethical Requirements	300
Other Issues	300
TRUSTEE Use Cases Conclusion Remarks.....	301
Part III: User-centred, Legal, and Socio-ethical requirements	302
Part III Executive Summary	302
Part III Introduction.....	302
User Experience Research: Why and How to Conduct It	302
Methodology	303
UX Research for User-centered Requirements Extraction	304

Approach to Legal and Socio-Ethical Requirements Definition	308
Outcomes and Analysis	312
TRUSTEE Personas and User Stories	313
TRUSTEE User-centred Requirements	314
TRUSTEE Legal Requirements	319
TRUSTEE Socio-Ethical Requirements	321
Part IV: Technical Requirements	326
Part IV Summary	326
Part IV Introduction	326
TRUSTEE Application Building Blocks Technical Requirements	327
Methodology for Collecting Technical Requirements	327
AI Models as a Service (AIMaaS)	330
Accountable Transactions Recorder (ATR)	336
Authentication Manager (AM)	341
Data Protection Impact Assessment (DPIA)	346
Homomorphic Enabled Data Fusion (HEDF)	349
Data Lake (DL)	358
One Stop Shop (OneSS)	360
SSI Homomorphic Enabled Framework (SSI-HE)	362
Security & Trust Manager (STM)	367
Knowledge Repository (KR)	375
Trustworthy AI Support Design Framework (TAI-SDF)	378
TRUSTEE Dashboard (DA)	381
Part V: TRUSTEE Architecture	392
Part V Summary	392
TRUSTEE Architecture Development	393
Architecture Framework: TOGAF	393
Recommended Modelling Language: ArchiMate	396
TOGAF and ArchiMate Synergy	400

Recommended Tool for Describing TRUSTEE Architecture: Archi.....	400
TRUSTEE Architecture Principles	401
TRUSTEE Architecture Description.....	403
Legend: ArchiMate Core Elements and Relationships	403
Strategy Views.....	405
Business Views	407
Application Views.....	413
Technology, Physical Views.....	502
Implementation Views: Federated Learning Case.....	503
Motivation Views	508
TRUSTEE Holistic Conceptual Architecture	510
Annex 1: Information Concerning the Processing of the Data Shared in the TRUSTEE Survey	512
Annex 2: Surveys	516
2.1 Pilot Leader	517
2.2 Pilot Member	520
Annex 3: Personas and User Stories	525
3.1 Energy Domain.....	525
Personas.....	525
User Stories	526
3.2 Healthcare Domain	530
Personas.....	530
User Stories	531
3.3 Education Domain.....	534
Personas.....	534
User Stories	536
3.4 Automotive Domain.....	538
Personas.....	538
User Stories	539

3.5 Space Domain	542
Personas	542
User Stories	544
3.6 Multi-Disciplinary Data Exchange	546
Personas	546
User Stories	548
All Domains	549
Personas	549
User Stories	550
Annex 4: User-Centred Requirements and Interconnection with Legal Requirements	552
Annex 5: Legal Requirements.....	579
5.1 Legal Requirements for TRUSTEE End-Users	579
5.2 Legal Requirements for TRUSTEE System Developers	601
Annex 6: Socio-Ethical Requirements	624
6.1 Socio-ethical Requirements for TRUSTEE System Developers.....	624
6.2 Socio-Ethical Requirements for TRUSTEE End-Users	636

INDEX OF FIGURES

Figure 1: D2.1 parts and how they feed each other	43
Figure 2 Picture from “The Global Risks Report 2023” – Source: World Economic Forum Global Risks Perception Survey 2022-2023.....	47
Figure 3: AI system lifecycle vs actors per each stage	57
Figure 4: Current standards, guidelines, and procedures to ensure AI systems auditability	58
Figure 5: Technical tools from Working Group implementing Trustworthy AI	59
Figure 6: Semantic view of the EOSC framework	72
Figure 7: Schematic of the application of HE.....	164
Figure 8: FHE Threshold operation	167
Figure 9: FHE Threshold configuration	168
Figure 10: FHE Threshold decryption process	168
Figure 11: Example of CKKS threshold technique execution using OpenFHE library	169
Figure 12: Average of CT size.....	172
Figure 13: Schematic of an operation between two ciphertexts.	176
Figure 14: Encryption times for different frameworks in CKKS	177
Figure 15: Decryption times for different frameworks in CKKS.....	177
Figure 16: Evaluation times for different frameworks in CKKS.....	177
Figure 17: Increase of the error depending on the number of sums on a TC in CKKS.....	178
Figure 18: SSI flow	182
Figure 19: Having form an iterative solver for the optimization problem in (1) , denoted as $g(x, y)$, a deep learning architecture is derived where each layer of this network corresponds to a unique iteration of the solver.	200
Figure 20: Centralized FL mechanism	204
Figure 21: A demonstration of data partition in horizontal FL. In this example the two clients contain five personal features. However, each client has data for different persons. The image was retrieved from study. .	211
Figure 22: The FedAVg algorithm. Note that the global model at t-th communication round is denoted by θ_t and the local model of k-th client is denoted by θ_k	212
Figure 23: An illustration of horizontal federated learning with personalization layers	214
Figure 24: The parameter tree created by cluster federated learning.....	216
Figure 25: Swarm Learning training steps ⁷⁶	218

Figure 26: Swarm Learning architecture.....	219
Figure 27: Different training strategies	221
Figure 28: A generic HE-Federated learning scheme.....	227
Figure 29: Integrated simulation framework.....	230
Figure 30: The architecture of SqueezeDet	233
Figure 31: Architecture of PointPillars.....	235
Figure 32: An example of object detection when the object detectors are LIDAR-based, Image-absed and Fusion-based	236
Figure 33: Popular types NNs used in SLAM systems	239
Figure 34: Superpoint architecture.....	240
Figure 35: Siamese training used in UnsuperPoint. The original image and the warped are forwarded through the network. Correspondences between the two images are used for training the model (Christiansen, Kragh, Brodskiy, & Karstoft, 2019).....	241
Figure 36: Overview of SuperGlue	243
Figure 37: Architecture of solution proposed by LoFTR	243
Figure 38: Architecture of DeepVCP	245
Figure 39: The workflow of a super-resolution lidar methodology. From a sparse 3D point cloud derived from a 16-channel cheap lidar sensor, we can reconstruct a denser point cloud similar to that produced from a 64-channel lidar based on the deep unrolling method	247
Figure 40: Relation between RAM, RA, CA, and the System [image from].....	252
Figure 41: ISO/IEC/IEEE 42010:2011—Architecture Description [image from v1.9 IIRA].....	253
Figure 42: ISO/IEC 30141 Reference model and views.....	254
Figure 43: Functional domains, crosscutting functions and system characteristics.....	256
Figure 44: IoT reference model	257
Figure 45: The general structure of CREATE-IoT 3D RAM	259
Figure 46: IT perspective layers	260
Figure 47: IDS-RAM	262
Figure 48: Connector flow in IDS-RAM ⁹⁵³	263
Figure 49: ISO/IEC CD 30141 (https://www.iso.org/standard/65695.html)	264
Figure 50: Key concepts of the EIRA and their relationships.....	266

Figure 51: The BDVA reference architecture (source: www.bdva.eu)	268
Figure 52: OpenDei reference architecture framework	271
Figure 53: X-Road ecosystem.....	274
Figure 54: Cross-border data exchange [NIIS – X-Road and eDelivery]	275
Figure 55: EU AI Framework	277
Figure 56: Stakeholders accessing data	292
Figure 57: Bottom-up approach contact tracing	293
Figure 58: Top-down approach contact tracing	293
Figure 59: Simple stakeholder environment in an educational setting	295
Figure 60: Complex stakeholder environment in an educational setting.....	295
Figure 61: VTT objectives for data exchanges	300
Figure 62: TRUSTEE Survey Personas mapping.....	313
Figure 63: Extracted User-centred requirements from Surveys.....	314
Figure 64: TRUSTEE Data Provider Persona	315
Figure 65: TRUSTEE Data Provider User Story	315
Figure 66: Finalized TRUSTEE User-centred requirements.....	318
Figure 67: TOGAF Architecture Development Cycle (image from)	394
Figure 68: Top-Level Hierarchy of ArchiMate Concepts (source: https://pubs.opengroup.org/architecture/ArchiMate3-doc/chap03.html)	397
Figure 69: ArchiMate Full Framework (source: https://pubs.opengroup.org/architecture/ArchiMate3-doc/chap03.html)	398
Figure 70: ArchiMate Notation overview (image from)	399
Figure 71: Archi tool IDE (image from)	401
Figure 72: TRUSTEE capabilities and resources	406
Figure 73: TRUSTEE stakeholders, drivers, goals, expected outcomes, courses of action, business capabilities and resources	407
Figure 74: TRUSTEE end-user groups, business services, interfaces, processes.....	408
Figure 75: TRUSTEE business capabilities to business processes	409
Figure 76: Data(set) activities in TRUSTEE	409
Figure 77: Data Provider business service to business processes to application services, processes, and data...	410

Figure 78: Model Provider business service to business processes to application services, processes, and data	411
Figure 79: Consumer using data - business service to business processes to application services, processes, and data	412
Figure 80: Consumer using model - business service to business processes to application services, processes, and data	412
Figure 81: Developer business service to business processes to application services, processes, and data	413
Figure 82: Application Building Blocks overall interaction diagram	415
Figure 83: TRUSTEE Data Provider Use Case Flow diagram	416
Figure 84: TRUSTEE Model Provider Use Case Flow diagram	417
Figure 85: Data Consumer Use Case sub-Flow diagram	418
Figure 86: Model Consumer Use Case sub-Flow diagram	418
Figure 87: Developer Use Case Flow diagram	419
Figure 88: User authentication and interaction with STM and PIA-CNIL components	420
Figure 89: DA decomposition diagram and external interfaces	421
Figure 90: DA-AM register sequence diagram	422
Figure 91: DA-AM login sequence diagram	422
Figure 92: DA request history of activity sequence diagram from ATR	423
Figure 93: DA HE-SSI-VERIFIER visualizing search operations sequence diagram	423
Figure 94: DA-DPIA sequence diagram	424
Figure 95: DA retrieve content from KR sequence diagram	424
Figure 96: DA-STM sequence diagram	425
Figure 97: DA-AIMaaS sequence diagram	425
Figure 98: DA-HEDF sequence diagram	426
Figure 99: TRUSTEE AM decomposition diagram and external interfaces	431
Figure 100: TRUSTEE AM onboarding (registering) new user on the TRUSTEE system	432
Figure 101: TRUSTEE AM authenticating (login) user on TRUSTEE's Dashboard	432
Figure 102: STM context	435
Figure 103: STM data flows	436
Figure 104: STM and Data Provider interactions while creating an agreement for a dataset	437

Figure 105: STM and Consumer interactions while using data and/or AI models	438
Figure 106: STM sequence diagram, validity check.....	439
Figure 107: STM sequence diagram, communication with DA and compliance checker	440
Figure 108: TRUSTEE DPIA decomposition diagram and external interfaces.	444
Figure 109: DPIA context and breakdown	444
Figure 110: TRUSTEE DPIA user creating, editing, reviewing, validating/signing/rejecting assessment and ABB recovering DPIA status.	445
Figure 111: Transaction Recording phase at ATR	449
Figure 112: ATR decomposition diagram and External interfaces.....	449
Figure 113: ATR record ABB transaction diagram (here ABB showcased is STM)	450
Figure 114: TRUSTEE SSI-HE Framework high level architecture	453
Figure 115: TRUSTEE SSI-HE Framework external interactions	454
Figure 116: TRUSTEE SSI-HE Framework operation flow.....	455
Figure 117: TAI-SDF	462
Figure 118: TAI-SDF architecture with inner modules and external interfaces.....	463
Figure 119: AIMaaS application view diagram	470
Figure 120: Sequence diagram of the use case of a Data Consumer using the AIMaaS.....	472
Figure 121: Sequence diagram of the use case of a Model Provider using the AIMaaS	473
Figure 122: DL ABB breakdown diagram	479
Figure 123: Sequence diagram: DL – AIMaaS (get model & store new runs).....	480
Figure 124: Sequence diagram: DL – SSI-HE (search attributes)	480
Figure 125: Sequence diagram: DL – SSI-HE (get attribute catalog).....	481
Figure 126: Sequence diagram: DL – HEDF (provide indexes & models).....	481
Figure 127: Sequence diagram: DL – HEDF (store new indexes)	482
Figure 128: Sequence diagram: DL – SSI-HE, obtaining public keys	482
Figure 129: Sequence diagram: DL – SSI-HE, providing public keys	482
Figure 130: Sequence diagram: DL – KR (search for previous results)	483
Figure 131: Sequence diagram: DL – Distributed Data Sources & External Connectors (store new data)	483
Figure 132: SDK external connectors ABB breakdown	489

Figure 133: HEDF ABB architecture	492
Figure 134: HEDF ABB diagram.....	493
Figure 135: HEDF Sequence diagram.....	494
Figure 136: Content gets stored to an IPFS node and the generated CID is sent to the USER. The CID is then stored in the blockchain.	499
Figure 137: Knowledge Repository decomposition diagram and external interfaces	499
Figure 138: Content stored at Knowledge Repository and ATR integration for transaction recording	500
Figure 139: TRUSTEE infrastructure serving the needs of the selected TRUSTEE Applications	503
Figure 140: FL case for automotive data domain: ABB components and connections	504
Figure 141: TRUSTEE Motivational layers view	509
Figure 142: TRUSTEE Motivation layers from Stakeholders to Goals	510
Figure 143: TRUSTEE conceptual architecture	511

INDEX OF TABLES

Table 0.1 D2.1 parts (high level chapters) mapping to WP2 objectives	42
Table 0.2 Semantic ontologies for the Health domain	65
Table 0.3 Semantic ontologies for the Education domain.....	66
Table 0.4 Semantic ontologies for the Space domain	67
Table 0.5 Semantic ontologies for the Automotive domain.....	68
Table 0.6 Semantic ontologies for the Energy domain.....	69
Table 0.7 Domain-independent semantic ontologies.....	70
Table 0.8 Summary of HE techniques.....	170
Table 0.9 Summary of the main FHE frameworks (17/01/23) [Part 1].	171
Table 0.10 Summary of the main FHE frameworks (17/01/23) [Part 2].....	171
Table 0.11 Operations implemented by the main FHE libraries.....	173
Table 0.12. Comparison of IdM systems.....	180
Table 0.13 Comparative study of FL frameworks (N/A indicates missing information).	205
Table 0.14 Requirements Description Template	329
Table 0.15 ArchiMate Core Elements	404
Table 0.16 ArchiMate Relationships.....	404
Table 0.17 Application Layer Elements (from ArchiMate® 3.1 Specification)	414

GLOSSARY OF TERMS

Abstraction - the technique of providing summarized or generalized descriptions of detailed and complex content. (TOGAF)

Access Control - means to ensure that access to assets is authorized and restricted based on business and security requirements. [ISO/IEC 27000: 2018]

Access Control Policy - the set of rules that define the conditions under which an access may take place [NISTIR 7316]

Accountability - principle that guarantees data controller's compliance with data protection principles, through the adoption of technical and internal measures (Article 5(2) GDPR). In the context of the AI, a set of mechanisms, practices and for the stewardship of personal and/or confidential data with which it [data organization] is entrusted in a cloud environment, for processing, storing, sharing, deleting and otherwise using data according to contractual and legal requirements from the time it is collected until when the data are destroyed (including onward transfer to and from third parties). Accountability involves committing to legal and ethical obligations, policies, procedures and mechanism, explaining (Felici, Loulours, Pearson 2013) (source: (Computational disciplines) IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. URL: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/eadv2_glossary.pdf).

Accountable - answerable for actions, decisions, and performance. [[ISO/IEC DIS 22989\(en\)](#). [Terms related to Trustworthiness](#)]

Actor - a person, organization, or system that has one or more roles that initiates or interacts with activities; for example, a sales representative who travels to visit customers. Actors may be internal or external to an organization. (TOGAF)

Admin – TAI-SDF user being able to add users and label them either a manager or data entry

Aggregated data - data combined or collected together in summary or other form such that the data cannot be identified with any individual [Legal Information Institute, Cornell University]

Air Gap - an interface between two systems at which (a) they are not connected physically and (b) any logical connection is not automated (i.e., data is transferred through the interface only manually, under human control) [[CNSSI 4009-2015](#) from [IETF RFC 4949 Ver 2](#)]

Algorithm - a formula, set of instructions, processes, rules or steps for performing a specific task

Anonymity - a condition in identification whereby an entity can be recognized as distinct, without sufficient identity information to establish a link to a known identity [[NISTIR 8053](#) from ISO/IEC 24760-1:2011]

Anonymization – process that removes the association between the identifying dataset and the data subject [[NISTIR 8053](#) from ISO/TS 25237:2008]. Anonymisation consists in preventing any identification of individuals from personal data. [[HLEG AI, Assessment List for Trustworthy AI \(ALTAI\)](#)] A process of removing personal identifiers, both direct and indirect, to achieve a stage in which a data subject cannot be re-identified by either the controller or another person, by using means reasonably likely to be used by them to re-identify an individual, whereby in the assessment of the “likelihood” the costs of identification, the time required for this and the available technology are all to be taken into account. (GDPR, Recital 26).

Anonymous data - information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable by a data controller or another person. Anonymisation is permanent and irreversible. Principles of data protection do not apply to anonymous information. [GDPR Recital 26]

Anticipatory Ethics - analysis of the standards for good or bad actions taken when designing, developing, deploying, or decommissioning emerging technologies. [(Ethics and Philosophy) IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems]

Application Programing Interface (API) - set of well-defined methods, functions, protocols, routines or commands which application software uses with facilities of programming languages to invoke services [ISO/TS 23029:2020]

Application Architecture provides a blueprint for the individual applications to be deployed, their interactions, and their relationships to the core business processes of the organization (TOGAF).

Application collaboration - represents an aggregate of two or more application internal active structure elements that work together to perform collective application behavior. (TOGAF)

Application Component - an encapsulation of application functionality aligned to implementation structure, which is modular and replaceable. It encapsulates its behavior and data, provides services, and makes them available through interfaces. (TOGAF)

Application event - represents an application state change. (TOGAF)

Application function - represents automated behavior that can be performed by an application component. (TOGAF)

Application interaction - represents a unit of collective application behavior performed by (a collaboration of) two or more application components. (TOGAF)

Application interface - represents a point of access where application services are made available to a user, another application component, or a node. (TOGAF)

Application process - represents a sequence of application behaviors that achieves a specific result. (TOGAF)

Application service - represents an explicitly defined exposed application behavior.

Architecture - fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and the principles of its design and evolution. [ISO/IEC/IEEE 42010]

Architecture Building Block (ABB) - a constituent of the architecture model that describes a single aspect of the overall model. An Architecture Building Block describes generic characteristics and functionalities. Architecture Building Blocks are used to describe reference architectures, solution architecture templates or solution architectures of specific solutions. (TOGAF) (see **Building Block**)

Architecture Domain - the architectural area being considered. The TOGAF framework has four primary architecture domains: business, data, application, and technology. Other domains may also be considered (e.g., security). (TOGAF)

Architecture Framework - conventions, principles, and practices for the description of architectures established within a specific domain of application and/or community of stakeholders. [ISO/IEC/IEEE 42010]

Architecture Principle - a qualitative statement of intent that should be met by the architecture. (TOGAF)

Architecture view - a work product expressing the architecture of a system from the perspective of specific system concerns. [[NIST SP 800-160 Vol. 1](#) from ISO/IEC/IEEE 42010]

Architecture Viewpoint - a specification of the conventions for a particular kind of architecture view. (TOGAF). A work product establishing the conventions for the construction, interpretation, and use of architecture views to frame specific system concerns. [[NIST SP 800-160 Vol. 1](#) from ISO/IEC/IEEE 42010]

Artificial intelligence (AI) – Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. The AI system can generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning approaches (including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning), machine reasoning (which includes, in particular, planning, scheduling, search, and optimization), logic- and knowledge-based approaches (including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems), statistical approaches, Bayesian estimation and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber- physical systems). [White Paper On Artificial Intelligence - A European approach to excellence and trust. European Commission. COM(2020) 65 final; [HLEG AI, Assessment List for Trustworthy AI \(ALTAI\)](#); Proposal of the EU Artificial Intelligence Act (Brussels, 21.4.2021 COM(2021) 206 final 2021/0106 (COD)), Article 2, point 1.]

Artificial Intelligence System (AI system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments. (DRAFT Compromise Amendments on the Draft Report Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9 0146/2021 – 2021/0106(COD))

Artificial Intelligence System Lifecycle – AI system lifecycle phases involve: i) ‘design, data and models’; which is a context-dependent sequence encompassing planning and design, data collection and processing, as well as model building; ii) ‘verification and validation’; iii) ‘deployment’; and iv) ‘operation and monitoring’. These phases often take place in an iterative manner and are not necessarily sequential. The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase. [[OECD AI Principles](#)]

- (a) as regards a controller with establishments in more than one Member State, the place of its central administration in the Union, unless the decisions on the purposes and means of the processing of personal data are taken in another establishment of the controller in the Union and the latter establishment has the power to have such decisions implemented, in which case the establishment having taken such decisions is to be considered to be the main establishment;
- (b) as regards a processor with establishments in more than one Member State, the place of its central administration in the Union, or, if the processor has no central administration in the Union, the establishment of the processor in the Union where the main processing activities in the context of the activities of an establishment of the processor take place to the extent that the processor is subject to specific obligations under the GDPR. [Article 4(16) GDPR]

Asset - anything that has value to the organization, its business operations and their continuity, including information resources that support the organization’s mission [ISO 27000]

Asymmetric cryptography - cryptography that uses two separate keys to exchange data, one to encrypt or digitally sign the data and one for decrypting the data or verifying the digital signature. Also known as public key cryptography. [[NIST SP 800-77 Rev. 1](#) under Asymmetric Cryptography]

Attack - an attempt to gain unauthorized access to system services, resources, or information, or an attempt to compromise system integrity.

Attack surface - the set of points on the boundary of a system, a system element, or an environment where an attacker can try to enter, cause an effect on, or extract data from, that system, system element, or environment. [[NIST SP 800-172](#) from [GAO-19-128](#)]

Attribute - inherent property or characteristic of an entity that can be distinguished quantitatively or qualitatively by human or automated means [ISO/IEC 25000:2014]

Audit - independent review and examination of records and activities to assess the adequacy of system controls, to ensure compliance with established policies and operational procedures and to recommend necessary changes in controls, policies or procedures [NISTIR-7298 rev 2]

Authentication - provision of assurance that a claimed characteristic of an entity is correct. (ISO/IEC 27000: 2018)

Authentication information - information used to establish the validity of a claimed identity.

Authorize - a decision to grant access, typically automated by evaluating a subject's attributes. [[NIST SP 800-63-3](#)]

Automatic/Automated/Automation - pertaining to a process or system that, under specified conditions, functions without human intervention. [[ISO/IEC DIS 22989\(en\). Terms related to Artificial Intelligence](#)].

Availability - the degree to which a system or component is operational and accessible when required for use [IEEE Glossary]. Property of being accessible and usable on demand by an authorised entity. [ISO/IEC DIS 22989(en). Terms related to Trustworthiness]

Bayesian network - probabilistic model that uses Bayesian inference for probability computations using a directed acyclic graph. [[ISO/IEC DIS 22989\(en\). Terms related to Artificial Intelligence](#)]

Bias - inclination of prejudice towards or against a person, object, or position. Bias can arise in many ways in AI systems. For example, in data-drive AI systems, such as those produced through machine learning, bias in data collection and training can result in an AI system demonstrating bias. In logic-based AI, such as rule-based systems, bias can arise due to how a knowledge engineer might view the rules that apply in a particular setting. Bias can also arise due to online learning and adaptation through interaction. It can also arise through personalisation whereby users are presented with recommendations of information feeds that are tailored to the user's tastes. It does not necessarily relate to human bias or human-driven data collection. It can arise, for example, through the limited contexts in which a system is used, in which case there is no opportunity to generalise it to other contexts. Bias can be good or bad, intentional or unintentional. In certain cases, bias can result in discriminatory and/or unfair outcomes, indicated in (HLEG AI, 2019) as unfair bias. [[HLEG AI, Ethics Guidelines for Trustworthy AI](#)]

Big Data - means extensive datasets - primarily in the data characteristics of volume, variety, velocity, and/or variability - that require a scalable technology for efficient storage, manipulation, management, and analysis. Big data is commonly used in many different ways, for example as the name of the scalable technology used to handle big data extensive datasets. [[ISO/IEC 20546:2019\(en\) Information technology - Big data - Overview and vocabulary](#)]

Big Data analytics – a subset of analytics focused on examining large volumes of data to uncover hidden patterns, correlations and other insights.

Binding corporate rules - means personal data protection policies which are adhered to by a controller or processor established on the territory of a Member State for transfers or a set of transfers of personal data to a controller or

processor in one or more third countries within a group of undertakings, or group of enterprises engaged in a joint economic activity. [Article 4(20) of the GDPR]

Biometric data — personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data. [Article 4(14) of the GDPR and [EU Artificial Intelligence Act](#)]

Blacklist — a security mechanism prohibiting the execution of those programs on a known malicious or undesired list of software. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Bodies governed by public law - bodies that have the following characteristics: (a) they are established for the specific purpose of meeting needs in the general interest, and do not have an industrial or commercial character; (b) they have legal personality; (c) they are financed, for the most part, by the state, regional or local authorities, or by other bodies governed by public law; or are subject to management supervision by those authorities or bodies; or have an administrative, managerial or supervisory board, more than half of whose members are appointed by the State, regional or local authorities, or by other bodies governed by public law. [Article 2(12) of the [Data Governance Act, COM \(2020\) 767 final](#)]

Building Block - a (potentially re-usable) component of enterprise capability that can be combined with other building blocks to deliver architectures and solutions. Building blocks can be defined at various levels of detail, depending on what stage of architecture development has been reached. For instance, at an early stage, a building block can simply consist of a name or an outline description. Later, a building block may be decomposed into multiple supporting building blocks and may be accompanied by a full specification. Building blocks can relate to "architectures" or "solutions". According to The Open Group, building blocks have several generic characteristics: see **Architecture Building Block, Solution Building Block**. According to The Open Group, building blocks have several generic characteristics: A building block is a package of functionality defined to meet the business needs across an organization, A building block has published interfaces to access the functionality, A building block may have multiple implementations but with different inter-dependent “building blocks.” A building block may interoperate with other, interdependent building blocks, A good building block considers implementation and usage and evolves to exploit technology and standards, It may be assembled from other building blocks, It may be a subassembly of other building blocks, Ideally, a building block is re-usable and replaceable, and well-specified.

Business Actor - represents a business entity that can perform behavior. (TOGAF)

Business Architecture - a representation of holistic, multi-dimensional business views of: capabilities, end-to-end value delivery, information, and organizational structure; and the relationships among these business views and strategies, products, policies, initiatives, and stakeholders. (TOGAF)

Business Capability - a particular ability that a business may possess or exchange to achieve a specific purpose. (TOGAF)

Business Function - delivers business capabilities closely aligned to an organization, but not necessarily explicitly governed by the organization. (TOGAF)

Business Governance - concerned with ensuring that the business processes and policies (and their operation) deliver the business outcomes and adhere to relevant business regulation. (TOGAF)

Business Interface - represents a point of access where a business service is made available to the environment. (TOGAF)

Business Model - a model describing the rationale for how an enterprise creates, delivers, and captures value. (TOGAF)

Business Process - represents a sequence of business behaviors that achieves a specific result such as a defined set of products or business services. (TOGAF)

Business Role - represents the responsibility for performing specific behavior, to which an actor can be assigned, or the part an actor plays in a particular action or event. (TOGAF)

Business Service - represents explicitly defined behavior that a business role, business actor, or business collaboration exposes to its environment. (TOGAF) Supports business capabilities through an explicitly defined interface and is explicitly governed by an organization. (TOGAF)

Capability - an ability that an organization, person, or system possesses. (TOGAF)

Capability - represents an ability that an active structure element, such as an organization, person, or system, possesses. (TOGAF)

Capability Architecture - a highly detailed description of the architectural approach to realize a particular solution or solution aspect. (TOGAF)

Certificate Authority (CA) - a trusted entity that issues and revokes public key certificates. [NIST SP 1800-16B under Certificate Authority from NISTIR 8149]

Ciphertext (CT)— the unintelligible and seeming random form of data that is produced by the cryptographic function of encryption. Ciphertext is produced by a symmetric algorithm when a data set is transformed by the encryption process using a selected key. Ciphertext can be converted back into its original form (i.e. plain text) by performing the decryption process using the same symmetric encryption algorithm and the key used during the encryption process. (Also known as a cryptogram.) [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Cloud - a network of remote servers on the internet that host, store, manage, and process applications and their data

Cloud Computing - paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand. [[ISO/IEC 20546:2019\(en\) Information technology - Big data - Overview and vocabulary.](#)]

Cloud environment - a cloud environment composed of hardware and software components used for aggregation of data coming from different sources (such as sensors) to store, analyze, transform and make data available to users.

Concern - an interest in a system relevant to one or more of its stakeholders. (TOGAF)

Confidentiality - the property that information is not made available or disclosed to unauthorized individuals, entities, or processes. The purpose of the confidentiality service is to protect information from unauthorized disclosure. [ITU-TX.800]

Consent - consent of the data subject means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her. [Article 4(11) of the GDPR]

Context of use - users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used [ISO/IEC 25000:2014]

Control - a measure that is modifying risk. Controls include any process, policy, device, practice, or other actions that modify risk. (ISO/IEC 27000: 2018)

Controller – means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of

such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law. [Article 4(7) of the GDPR]

Cross-border processing – means either processing of personal data which takes place in the context of the activities of an organisation in more than one Member State where that organisation is established in more than one Member State; or processing of personal data which takes place in the context of the activities of an organisation's single establishment but where that processing substantially affects or is likely to substantially affect data subjects in more than one Member State. [GDPR, Article 4(23)]

Cyberattack — any attempt to violate the security perimeter of a logical environment. An attack can focus on gathering information, damaging business processes, exploiting flaws, monitoring targets, interrupting business tasks, extracting value, causing damage to logical or physical assets or using system resources to support attacks against other targets. Cyberattacks can be initiated through the exploitation of a vulnerability in a publicly exposed service, through tricking a user into opening an infectious attachment or even causing automated installation of exploitation tools through innocent website visits. (Also known as a drive-by download.) [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Cybersecurity — the efforts to design, implement, and maintain security for an organization's network, which is connected to the Internet. It is a combination of logical/technical-, physical- and personnel-focused countermeasures, safeguards and security controls. An organization's cybersecurity should be defined in a security policy, verified through evaluation techniques (such as vulnerability assessment and penetration testing) and revised, updated and improved over time as the organization evolves and as new threats are discovered [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Data - a representation of facts, concepts, or instructions in a manner suitable for communication, interpretation, or processing by humans or by automatic means [IEEE Glossary]. A digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording. [DGA, [Article 2\(1\)](#)]

Data aggregation - the process of selecting data from different sources and aggregates them to create new knowledge value. Aggregation, along with K-anonymity technique, aims to prevent a data subject from being singled out by grouping them with, at least, k other individuals. From the data protection perspective on anonymization, data aggregation is one of the generalization techniques (generalization is one of the families of anonymization techniques). [Opinion 05/2014 on Anonymisation Techniques, 0829/14/EN, WP216]

Data analysis - process of transforming raw data into usable information, often presented in the form of a published analytical article, in order to add value to the statistical output. [OECD]

Data annotation - process of attaching a set of descriptive information to data without any change to that data. The descriptive information can take the form of metadata, labels and anchors. [[ISO/IEC DIS 22989\(en\). Terms related to Machine Learning](#)]

Data Architecture - a description of the structure and interaction of the enterprise's major types and sources of data, logical data assets, physical data assets, and data management resources. (TOGAF)

Data augmentation - process of creating new data samples by manipulating the original data. [[ISO/IEC DIS 22989\(en\). Terms related to Machine Learning](#)]

Data availability - the fact that data is accessible and services are operational. (ENISA)

Data breach — the occurrence of disclosure of confidential information, access to confidential information, destruction of data assets, or abusive use of a private IT environment. Generally, a data breach results in internal data being made accessible to external entities without authorization. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]. When it comes to personal data breach: it means a breach of security leading to the accidental or

unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed [Article 4(12) of the GDPR]

Data cleaning - the process of removing wrong data by using, for instance, Data anomaly detection, remove useless data for the application.

Data collection – (in the context of IoT) the process used by IoT application to run business logic might need to collect data from different sources that might be the sensors or the databases where datasets are stored.

Data concerning health - means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status [Article 4(15) of the GDPR].

Data Confidentiality - the property that information is not made available or disclosed to unauthorized individuals, entities, or processes. (ISO/IEC PDTR 13335-1). Confidentiality in a general sense refers to the duty not to share information with persons who are not qualified to receive that information (see Article 5(f) of Regulation (EU) 2016/679) [[EDPS Data protection Glossary](#)].

Data controller – see Controller

Data discovery - the process that will enable the application to access the IoT data without the need to know the actual source of data, sensor description, or location. It is enabled by the data abstraction techniques that hide the device heterogeneity of the applications.

Data federation is an aspect of data virtualization where the data stored in a heterogeneous set of autonomous data stores are made accessible to data consumers as one integrated data store by using on-demand data integration.

Data Entry – TAI-SDF user who fills out questionnaires on the TAI-SDF UI.

Data filtering - the process of removing duplicate data.

Data governance - a set of activities aimed to design, implement and monitor a strategic plan for data management.

Data Governance Act (DGA) - regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724.

Data holder – legal person, including public sector bodies and international organisations, or a natural person who is not a data subject with respect to the specific data in question, which, in accordance with applicable law, has the right to grant access to or to share certain personal data or non-personal data [Article 2(8)DGA]; means a legal or natural person, who has accessed data from the connected product or has generated data during the provision of a related service and who has the contractually agreed right to use such data, and the obligation, in accordance with this Regulation, applicable Union law or national legislation implementing Union law to make available certain data to the user or a data recipient [Data Act, Article 2(6), version after amendments of the EP, available at: [Texts adopted - Data Act - Tuesday, 14 March 2023 \(europa.eu\)](#)]

Data ingestion - the process of absorbing the bulk of data specified by the applications without losing the value of the original raw data.

Data integration - the process of combining different datasets, it deals with the data format heterogeneity. For example, combine a dataset using a certain database standard with another dataset using other database standards requires a syntactic and semantic interpretation of the datasets to be able to integrate them correctly.

Data integrity - the property that data has not been altered or destroyed in an unauthorized manner. [b-ITU-T X.800; ISO/IEC PDTR 13335-1]

Data intermediation service - a service which aims to establish commercial relationships for the purposes of data sharing between an undetermined number of data subjects and data holders on the one hand and data users on the other, through technical, legal or other means, including for the purpose of exercising the rights of data subjects in relation to personal data, excluding at least the following: (a) services that obtain data from data holders and aggregate, enrich or transform the data for the purpose of adding substantial value to it and license the use of the resulting data to data users, without establishing a commercial relationship between data holders and data users; (b) services that focus on the intermediation of copyright-protected content; (c) services that are exclusively used by one data holder in order to enable the use of the data held by that data holder, or that are used by multiple legal persons in a closed group, including supplier or customer relationships or collaborations established by contract, in particular those that have as a main objective to ensure the functionalities of objects and devices connected to the Internet of Things; (d) data sharing services offered by public sector bodies that do not aim to establish commercial relationships; [DGA, Article 2(11)]

Data lifecycle - includes processing and management activities conducted on data from its creation to its use and disposal.

Data management - the activities of defining, creating, storing, maintaining and providing access to data and associated processes in one or more information systems. [b-ISO/IEC TR 10032:2003]

Data masking - the process of making the data not possible to link with its owner. Masking is a broad term which refers to functions that when applied to data, they hide their true value. The most prominent examples are encryption and hashing but as the term is rather broad, it also covers additional techniques, some of which will be discussed within this section. The main usability of masking with regards to data protection principles is integrity and confidentiality (security) and depending on the technique or the context of the processing operation it can also include accountability and purpose limitation [[ENISA, Data Protection Engineering](#)]. Techniques as anonymization and pseudonymization are possible to enable the privacy of data during its lifecycle.

Data Mining — the activity of analyzing and/or searching through data in order to find items of relevance, significance, or value. The results of data mining are known as meta-data. Data mining can be a discovery of individual important data items, a summary or overview of numerous data items or consolidation or clarification of a collection of data items. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Data object- represents data structured for automated processing.

Data origin authentication - the corroboration that the source of data received is as claimed. [b-ITU-T X.800]

Data processing - the ability to handle data as input and apply different treatments that might modify the data or combine it without modifying it with other data in order to produce an output that is useful for a given application or service in the data lifecycle. In terms of the GDPR: *Any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction* [Article 4(2) of the GDPR]. Related definition – **restriction of processing** (see below).

Data processing and management (DPM) - the combination of all activities either directly performed on or indirectly influencing data. [b-FG-DPM TS D0.1]

Data processor - see **Processor**.

Data protection - means the implementation of appropriate administrative, legal, technical or physical means to guard against unauthorized intentional or accidental disclosure, modification, or destruction of data.

Data provider - organization which produces or collects and further shares data or metadata. [OECD]

Data recipient – in the context of personal data protection: means a natural or legal person, public authority, agency or another body, to which the personal data are disclosed, whether a third party or not. However, public authorities

which may receive personal data in the framework of a particular inquiry in accordance with Union or Member State law shall not be regarded as recipients; the processing of those data by those public authorities shall be in compliance with the applicable data protection rules according to the purposes of the processing. [Article 4(9) of the GDPR]. In the context of Data Act: means a legal or natural person other than the user of a connected product or related service, to whom a data holder makes available data accessed from a connected product or generated during the provision of a related service following an explicit request by the user or in accordance with a legal obligation under Union law or national legislation implementing Union law [Data Act, Article 2(7), version after amendments of the EP, available at: [Texts adopted - Data Act - Tuesday, 14 March 2023 \(europa.eu\)](#)]

Data sampling - process to select a subset of data samples intended to present patterns and trends similar to that of the larger dataset being analysed. [[ISO/IEC DIS 22989\(en\). Terms related to Machine Learning](#)]

Data securing - the process to control the access to the data, to ensure its confidentiality and integrity.

Data sharing - data sharing is not only about communication technology interoperability. It is the process of combining data from different sources (devices, databases...) and this is possible only by identifying and applying the business intelligence related rules to allow data sharing, and the communication and metadata standards to enable the effective use of data owned by different players and stored and processed by different technologies. According to the DGA, 'data sharing' means the provision of data by a data subject or a data holder to a data user for the purpose of the joint or individual use of such data, based on voluntary agreements or Union or national law, directly or through an intermediary, for example under open or commercial licences subject to a fee or free of charge. [DGA, Article 2(10)]

Data source - specific data set, metadata set, database or metadata repository from where data or metadata are available. [OECD]

Data storage - the process to accumulate data for future processing and use. The duration of storage depends on the application and the security, privacy, and governance rules.

Data subject - natural person whose personal data is processed by a data controller or processor. An identified or identifiable natural person. [GDPR, Article 4(1)]

Data Theft - the act of intentionally stealing data. Data theft can occur via data loss (physical theft) or data leakage (logical theft) event. Data loss occurs when a storage device is lost or stolen. **Data leakage** occurs when copies of data are possessed by unauthorized entities. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Data transfer – any transfer of data which are undergoing processing or intended for processing after to a third country or to an international organisation [Article 44 GDPR]. Transfers of personal data when the recipient is located in a country outside the EU/EEA is subject to Chapter V of the Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation).

Data trust - includes actions taken to safeguard security, privacy, and quality of data and enhance data trustworthiness.

Data use - after the data visualization and analysis, the data use is the process that application run to either understand the monitored phenomena by analyzing the past and monitor the present or predict the future and build new decisions.

Data user - a natural or legal person who has lawful access to certain personal or non-personal data and has the right, including under the GDPR in the case of personal data, to use that data for commercial or non-commercial purposes. [DGA, Article 2(9)]

Data validation - the process to check the quality, the correctness, and trustworthiness of data.

Data visualization - the process of displaying the data to be analyzed for the application of business intelligence purpose.

Database - a collection of interrelated data stored together in one or more computerized files. [IEEE Glossary]

Dataset - any organised collection of data. [OECD]

Distributed Denial of Service (DDoS) attack — an attack that attempts to block access to and use of a resource. It is a violation of availability. DDOS (or DDoS) is a variation of the DoS attack (see DOS) and can include flooding attacks, connection exhaustion, and resource demand. The distinction of DDOS from DOS is that the attack traffic may originate from numerous sources or is reflected or bounced off of numerous intermediary systems. The purpose of a DDoS attack is to significantly amplify the level of the attack beyond that which can be generated by a single attack system in order to overload larger and more protected victims. DDoS attacks are often waged using botnets. (See botnet.) [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

De-identification - removal of personally identifying information in order to protect personal privacy. **De-identification** is the process for protecting personal information; this refers to the use of randomization, masking, anonymity, and concealment to prevent personal information leakage, by making it impossible for others to identify the personal data of specific persons [...] As de-identified data is no longer treated as personal data, it should be impossible to acquire the identities of specific persons from this data. That is, de-identification is only considered successful if it is impossible to identify individuals even with tracking or comparative techniques. (Hsieh-Hong Huang, Jian-Wei Lin and Chia-Hsuan Lin ‘Data Re-Identification—A Case of Retrieving Masked Data from Electronic Toll Collection’ (Symmetry 2019) 11(4))

Design - the process of defining the architecture, components, interfaces, and other characteristics of a system or component [IEEE Glossary]

Developer - individual or organization that performs development activities (including requirements analysis, design, testing through acceptance) during the system or software life-cycle process [: ISO/IEC 25000:201]

Device - equipment with the mandatory capabilities of communication and the optional capabilities of sensing, actuation, data capture, data storage and data processing. Hardware component responsible for all explicit (i.e., handled by presentation devices) and implicit (input (sensors) and output (actuator)) interaction with the end-user.

Differential Privacy - a mechanism for publicly sharing details regarding a dataset by defining the patterns of groups within the dataset while keeping information about individuals inside the dataset private.

Digital signature - data appended to, or a cryptographic transformation (see cryptography) of a data unit that allows a recipient of the data unit to prove the source and integrity of the data unit and protect against forgery, e.g., by the recipient. [b-ITU-T X.800]

Electronic health record (EHR) - personal data created, developed, maintained, and/or provided by clinicians, providers, and allied health providers in direct patient care. An electronic application containing health information about individuals that is used by clinicians, providers, and allied health professionals to provide direct care for the individuals.

End-to-end security – approach to security where data traveling between clients and servers’ end-points are uninterruptedly protected, even where untrusted intermediaries entities or communication channels are required.

End-user - a person / entity that consumes or utilizes a product or service.

End-user environment – an environment in which the end-user interactions with the system occur. It includes specific devices, software and user interfaces directly available to the people that will use the system.

Enterprise - means a natural or legal person engaged in an economic activity, irrespective of its legal form, including partnerships or associations regularly engaged in an economic activity. [Article 4(18) of the GDPR]

Error - discrepancy between a computed, observed, or measured value or condition and the true, specified, or theoretically correct value or condition. (Source: ISO/IEC 15026-1: 2019)

European Health Data Space (EHDS) - the creation of a European Data Space is one of the priorities of the Commission 2019-2025, including the health sector. In the Communication on “A European strategy for data” it states that the Commission will support the establishment of nine common European data spaces with one of them being the Common European health data space, which is essential for advances in preventing, detecting and curing diseases as well as for informed, evidence-based decisions to improve the accessibility, effectiveness and sustainability of the healthcare systems. [European Commission]

Exception - an event that causes suspension of normal program execution, or an indication that an operation request was not performed successfully. (Source: ISO/IEC/IEEE 24765: 2017)

FAIR Data principles - the FAIR Data Principles are a set of guiding principles in order to make data findable, accessible, interoperable and reusable.

Fault isolation - the ability of a subsystem to prevent a fault within the subsystem from causing consequential faults in other subsystems. (Source: ISO/IEC/IEEE 24765: 2017)

Federated Learning - is a machine learning setting where many clients (e.g. mobile devices or whole organizations) collaboratively train a model under the orchestration of a central server (e.g. service provider), while keeping the training data decentralized.

Federation stems from the Latin words *foedus* and *foederis*, which both mean treaty, agreement, and contract. In most cases, if the term federation is used, it refers to combining autonomously operating objects.

Filing system – means any structured set of personal data which are accessible according to specific criteria, whether centralized, decentralized or dispersed on a functional or geographical basis. [Article 4(6) GDPR]

Firewall — a security tool, which may be a hardware or software solution that is used to filter network traffic. A firewall is based on an implicit deny stance where all traffic is blocked by default. Rules, filters or ACLs can be defined to indicate which traffic is allowed to cross the firewall. Advanced firewalls can make allow/deny decisions based on user authentication, protocol, header values and even payload contents. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Formal Caregiver - caregiver that is paid for being involved in service provision (e.g., service personnel working for service providers) or are otherwise involved professionally (e.g., skilled volunteers with social commitment).

Functional requirement - a functionality or service that the system has to offer

Governance - the discipline of monitoring, managing, and steering a business (or IS/IT landscape) to deliver the business outcome required. (TOGAF)

Green Energy – energy generated from natural resources, sun, wind, and water namely renewable energy resources.

Homomorphic Encryption (Homomorphic Cryptography) – encryption schemes allowing certain mathematical operations to be performed directly on ciphertexts, without prior decryption. Homomorphic encryption can be a powerful tool for leveraging multi-party computations, by enabling a participant to compute functions on values while keeping the values hidden.

Health data – see “Data concerning health”.

Identifier - the name, address, label, or distinguishing index of an object in a computer program. [IEEE Glossary]

Information - any communication or representation of facts, data, or opinions, in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audio-visual forms. (TOGAF)

Information Society Service - means any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services. For the purposes of this definition: (i) 'at a distance' means that the service is provided without the parties being simultaneously present; (ii) 'by electronic means' means that the service is sent initially and received at its destination by means of electronic equipment for the processing (including digital compression) and storage of data, and entirely transmitted, conveyed and received by wire, by radio, by optical means or by other electromagnetic means; (iii) 'at the individual request of a recipient of services' means that the service is provided through the transmission of data on individual request. [Article 4(25) of the GDPR; Article 1(1)(b) of Directive (EU) 2015/1435 of the European Parliament and the Council]

Integrity - the property that data has not been altered in an unauthorized manner. [ITU-T H.235] The property that data is authentic and has not been maliciously or accidentally altered during processing, storage or transmission. [Article 29 Data Protection Working Party, Opinion 5/2012 on Cloud Computing (WP 196 2012)]

Interface - 1) A shared boundary across which information is passed. (2) A hardware or software component that connects two or more other components for the purpose of passing information from one to the other. [IEEE Glossary]

International organization - an organisation and its subordinate bodies governed by public international law, or any other body which is set up by, or on the basis of, an agreement between two or more countries. [Article 4(26) of the GDPR]

Internet of Things (IoT) - describes the coordination of multiple machines, devices, and appliances connected to the Internet through multiple networks. These devices include everyday objects such as tablets and consumer electronics, and other machines such as vehicles, monitors and sensors equipped with communication capabilities that allow them to send and receive data. [GSMA]

Interoperability - the ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged. [b-ITU-T Y.101]

Legal Requirement - a responsibility placed on an entity, including statutory or regulatory obligations which an entity is required to fulfil, to ensure legal compliance of its actions.

Lifecycle - states involved in the management of an asset; the evolution of a system, product, service, project, or other human-made entity from conception through retirement. (Sources: ISO/IEC 12207: 2017; ISO/IEC 27034: 2011)

Malware (malicious software) — any code written for the specific purpose of causing harm, disclosing information or otherwise violating the security or stability of a system. Malware includes a wide range of types of malicious programs including virus, worm, Trojan horse, logic bomb, backdoor, Remote Access Trojan (RAT), rootkit, ransomware, and spyware/adware. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Metadata - data collected on any activity of a natural or legal person for the purposes of the provision of a data sharing service, including the date, time and geolocation data, duration of activity, connections to other natural or legal persons established by the person who uses the service. [DGA COM (2020) 767 FINAL]

Metamodel - a model that describes how and with what the architecture will be described in a structured way. (TOGAF)

Method - a defined, repeatable approach to address a particular type of problem. (TOGAF)

Modelling - a technique through construction of models which enables a subject to be represented in a form that enables reasoning, insight, and clarity concerning the essence of the subject matter. (TOGAF)

Multi-disciplinary data – data from different science/study fields.

Multi – Modal data – different modifications of data or different forms and origins of data.

Network – computer or data network using a wired, wireless or both types of connections for the purpose of transmitting, receiving and exchanging data, voice and video traffic

Non-disclosure agreement – a legal contract between two parties or among more parties stating the use and sharing of data among the parties for certain purposes but restricts access to.

Non-functional requirements – complementary to functional requirements, which describes what a system is supposed to do, non-functional requirements describe “how a system is supposed to be”.

Non-personal data - data other than personal data as defined in point (1) of Article 4 of Regulation (EU) 2016/679. [Article 2(3) of the DGA COM (2020) 767 FINAL]

Non-repudiation is the ability to prevent entities later falsely denying that they performed (or did not perform) an action. Non-repudiation is concerned with establishing evidence that can later be used to counter false claims. Recommendation ITU-T X.800 describes two forms of non-repudiation service: non-repudiation with proof of delivery, which is used to counter false denial by a recipient that the data has been received, and nonrepudiation with proof of origin, which is used to counter false denial by an originator that the data has been sent.

Ontology – a set of concepts and categories in a domain that includes its properties and the relations among them.

Open data - open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike. [Open Data Handbook]

OWASP (Open Web Application Security Project) — an Internet community focused on understanding web technologies and exploitations. Their goal is to help anyone with a website improve the security of their site through defensive programming, design, and configuration. Their approach includes understanding attacks in order to know how to defend against them. OWASP offers numerous tools and utilities related to website vulnerability evaluation and discovery as well as a significant amount of training and reference material related to all things web security. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Persona - a fictional characterization of a user.

Pattern - a technique for putting building blocks into context; for example, to describe a re-usable solution to a problem. (TOGAF)

Personal data - means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person [Regulation (EU) 2016/679 (General Data Protection Regulation), Article 4 (1)]. Data which relates to a living individual who can be identified: a) from those data; or b) from those data and other information which is in possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual. [b-PAS 185:2017]

Personal data breach - means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed. [Regulation (EU) 2016/679 (General Data Protection Regulation), Article 4 (12)].

Personalization - the quality of incorporating knowledge about the characteristics, capabilities, and preferences of each human user when performing actions in the virtual realm. As a result, the differences between individuals can

be accommodated adequately. Personalization can be seen as a special case of Context-awareness since the data representing the knowledge about characteristics, capabilities, and preferences of each human user are often classified as a subset of Context Personal data “‘Personal data’ shall mean any information relating to an identified or identifiable natural person (‘data subject’); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or one or more factors specific to his physical, physiological, mental, economic, cultural or social identity,” Art.2 (a) of Directive 95/46/EC. Within this deliverable personal data is the terminology used for legal considerations. See also Personally Identifiable Information.

Personally Identifiable Information (PII) - personally Identifiable Information is defined as any information about an individual maintained by an [entity], including any information that can be used to distinguish or trace an individual’s identity, such as name, social security number, date and place of birth, and any other information that is linked or linkable to an individual([NIST10] p. 2-1). PII is a widely used term for personal data in the domain of information security. Within this document, PII is used in relation to information security

Platform - a platform is a group of technologies that are used as a base upon which other software is run. It typically includes hardware architecture, an Operative System and runtime libraries.

Privacy - the right of individuals to control or influence what information related to them may be collected and stored and by whom and to whom that information may be disclosed. [b-ITU-T X.800]

Privacy by default - according to Regulation 679/2016 the controller shall implement mechanisms for ensuring that, by default, only those personal data are processed which are necessary for each specific purpose of the processing and are especially not collected or retained beyond the minimum necessary for those purposes, both in terms of the amount of the data and the time of their storage. In particular, those mechanisms shall ensure that by default, personal data are not made accessible to an indefinite number of individuals¹.

Privacy by design - a methodology according to which, taking into account state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of persons posed by the processing, at the time of the determination of the means for processing, at the time of the processing itself and during the retention period, the implementation of appropriate technical and organizational measures, such as pseudonymization, which are designed to implement data-protection principles, such as data minimization in an effective manner and the integration of the necessary safeguards into the processing in order to meet regulatory requirements and to protect the rights of data subjects. [adapted from GDPR, Article 25(1)]

Privacy Impact Assessment (PIA) – in other words Data Protection Impact Assessment (DPIA) – describes a process designed to identify risks arising out of the processing of personal data and to minimise these risks as far and as early as possible. DPIAs are important tools for negating risk, and for demonstrating compliance with the GDPR. Required by Article 35 of the GDPR where data processing “is likely to result in a high risk to the rights and freedoms of natural persons”.

Processed data - data that have been transformed from raw data or an earlier data stage into a more refined stage by data cleaning, sorting, linking, verifying and similar operations. [b-ISO 5127:2017]

- (a) processing of personal data which takes place in the context of the activities of establishments in more than one Member State of a controller or processor in the Union where the controller or processor is established in more than one Member State; or

¹ Article 34(2), Regulation 679/2016, adopted on 4 May 2016, in <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&qid=1465303957140&from=EN>

- (b) processing of personal data which takes place in the context of the activities of a single establishment of a controller or processor in the Union but which substantially affects or is likely to substantially affect data subjects in more than one Member State. [Article 4(23) of the GDPR]

Processor - natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller. [Article 4(8) of the GDPR]

Profiling - any form of automated processing of personal data consisting of using data to evaluate certain personal aspects relating to a natural person, in particular, to analyze and predict aspects concerning that person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements. [Article 4(4) GDPR]².

Protocol – a set of communication rules for hardware or software or both that define how the entities are to communicate with each other (e.g., HTTP protocol)

Prototype (vs. Pilot)–while both approaches are intended to test and verify a system, a pilot generally intends to test the full production system against a specific subset of the end users. In contrast, the prototype may be focused on validating and learning from specific system aspects, implying that the prototype may not be part of the production version of the system.

Pseudonymization - the processing of personal data in such a manner that the data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure non-attribution to an identified or identifiable natural person [Article 4(5) of the GDPR]

Raw data - data that has not been processed for use. [b-PAS 185:2017]

Re-use - the use by natural or legal persons of data held by public sector bodies, for commercial or non-commercial purposes other than the initial purpose within the public task for which the data were produced, except for the exchange of data between public sector bodies purely in pursuit of their public tasks [DGA, Article 2(2)]

Reasoning - reasoning is defined as using context for facilitating context-awareness (e.g., inferring and adding new facts to context), personalization (e.g., deducing users' situational preferences), reactivity (e.g., learning from behavior analysis), and pro-activity (e.g., analysis of correspondence between situations and needs) in smart environments.

Reference Model (RM) - an abstract framework for understanding significant relationships among the entities of [an] environment, and for the development of consistent standards or specifications supporting that environment. (TOGAF)

Representative - means a natural or legal person established in the Union who, designated by the controller or processor in writing pursuant to Article 27 GDPR, represents the controller or processor with regard to their respective obligations under the GDPR. [Article 4(17) GDPR].

Requirement - need or expectation that is stated, generally implied, or obligatory. [b-ISO 8000-2]

Requirements elicitation - process through which the acquirer and the suppliers of a system discover, review, articulate, understand, and document the requirements on the system and the life cycle processes [adapted from ISO/IEC/IEEE 24765:2010]

² Recital 71 and Article 22, Regulation 679/2016, adopted on 4 May 2016, in <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&qid=1465303957140&from=EN>.

Requirements engineering - interdisciplinary function that mediates between the domains of the acquirer and supplier to establish and maintain the requirements to be met by the system, software or service of interest [ISO/IEC/IEEE 29148:2018]

Requirements specification - structured collection of the requirements (functions, performance, design constraints, and attributes) of the software and its external interfaces [adapted from IEEE Std 1012:200]

Resource - represents an asset owned or controlled by an individual or organization. (TOGAF)

Restriction of processing - means the marking of stored personal data with the aim of limiting their processing in the future [Article 4(3) GDPR].

Risk – means a combination of the probability of occurrence of harm and the severity of that harm. [ISO13485:2016]

Risk Assessment - the process of identifying risks to organizational operations (including mission, functions, image, reputation), organizational assets, individuals, other organizations, and the Nation resulting from the operation of an information system. Part of risk management incorporates threat and vulnerability analyses and considers mitigations provided by security controls planned or in place. Synonymous with risk analysis. [CNSSI-4009]

Risk Effect of uncertainty on objectives, a positive or negative deviation from what is expected. [ISO 31000]

Risk Management — the process of performing a risk assessment and evaluating the responses to risk in order to mitigate or otherwise handle the identified risks. Countermeasures, safeguards, or security controls are to be selected that may eliminate or reduce risk, assign or transfer risk to others (i.e., outsourcing or buying insurance) or avoid and deter risk. The goal is to reduce risk down to an acceptable or tolerable level. [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Role Based Access Control (RBAC) – a.k.a. authorization.

Scenario – a single path that is comprised of distinct steps to accomplish a goal.

Security - the term "security" is used in the sense of minimizing the vulnerabilities of assets and resources. An asset is anything of value. A vulnerability is any weakness that could be exploited to violate a system or the information it contains. A threat is a potential violation of security. [ITU-TX.800]

Security by design – consider security requirements as early as possible in the system design and development and continuously integrate and perform relevant tests [Guidelines 4/2019 on Article 25 Data Protection by Design and by Default]

Security Gateway - means a set of control mechanisms between two or more networks having different trust levels which filter and log traffic passing, or attempting to pass, between networks, and the associated administrative and management servers.

Security Policy - definition of the objectives and constraints for the security program.

Security requirements – OWASP defines security requirements in its quick reference guide as “a set of design and functional requirements that help ensure the software is built and deployed in a secure manner.”

Security Testing - process to determine that an information system protects data and maintains functionality as intended [CNSSI-4009(4)]

Seismic events –quake events/cases, such as earthquakes, moonquakes, etc.

Sensitive data - any information that is protected against unwarranted disclosure. Protection of data may be required for legal or ethical reasons, for issues pertaining to personal privacy, or for proprietary considerations. [CSC]

Sensitive Personal Data / Special category of personal data – Personal data which are, by their nature, particularly sensitive in relation to fundamental rights and freedoms merit specific protection as the context of their processing could create significant risks to the fundamental rights and freedoms. In the GDPR sensitive personal data are considered as “special categories of personal data”. The following personal data is considered ‘special category of personal data’ and is subject to specific processing conditions stemming in particular from the GDPR: personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data processed for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation. [Article 4(13), 4(14), 4(15), 9(1) and Recitals 51 to 56 of the GDPR, https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en]

Sensor - a device that realizes a sensing channel by capturing info from the physical world (mostly in terms of measuring a quantity) and providing it in the form of data as contextual info within the virtual realm. Sensors are devices capturing signals from the caring home: the environment and the users. Some are mounted at the home ant its furniture, while others are wearable.)

Service – a) A repeatable activity; a discrete behavior that a building block may be requested or otherwise triggered to perform. b) An element of behavior that provides specific functionality in response to requests from actors or other services. (TOGAF)

Software development life cycle (SDLC) - the period of time that begins with the decision to develop a software product and ends when the software is delivered. This cycle typically includes a requirements phase, design phase, implementation phase, test phase, and some- times, installation and checkout phase

Software requirement - see Requirement

Software requirements elicitation - see Requirements elicitation

Software requirements engineering - see Requirements engineering

Software requirements specification - see Requirements specification

Software system - a set of information resources organized for the collection, processing, maintenance, use, and sharing of information.

Software. all or part of the programs that process or support the processing of digital information. (Source: ISO/IEC 12207: 2017)

Solution Building Block (SBB) - solution Building Block (SBB) is related to the Solutions Continuum. SBBs are components used to fulfil the requirements of ABBs, which will direct and guide SBB development. A candidate solution which conforms to the specification of an Architecture Building Block (see **Architecture Building Block**).

Stakeholders [b-ISO GUIDE 73]: person or organization that can affect, be affected by, or perceive themselves to be affected by a decision or activity

Standard – a consistent protocol that are universally understood and accepted

Strong Encryption means the use of encryption technologies with minimum key lengths of 256-bits for symmetric encryption and 1024-bits for asymmetric encryption whose strength provides reasonable assurance that it shall protect the encrypted information from unauthorized access and is adequate to protect the confidentiality and privacy of the encrypted information, and which incorporates a documented policy for the management of the encryption keys and associated processes adequate to protect the confidentiality and privacy of the keys and passwords used as inputs to the encryption algorithm. Strong Encryption includes, but is not limited to: SSL

v3.0+/TLS v1.0+, Point to Point Tunneling Protocol (PPTP), AES 256, FIPS 140-2 (United States government only), RSA 1024 bit, SHA1/SHA2/SHA3, Internet Protocol Security (IPSEC), SFTP, SSH, Vormetric v4, or WPA2.

Synthetic data - is artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data. This means that synthetic data and original data should deliver very similar results when undergoing the same statistical analysis. The degree to which synthetic data is an accurate proxy for the original data is a measure of the utility of the method and the model. (https://edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en). The generation process, also called *synthesis*, can be performed using different techniques, such as decision trees, or deep learning algorithms. Synthetic data can be classified with respect to the type of the original data: the first type employs real datasets, the second employs knowledge gathered by the analysts instead, and the third type is a combination of these two.

System - a combination of interacting elements organized to achieve one or more stated purposes. The interacting elements that compose a system include hardware, software, data, humans, processes, procedures, facilities, materials, and naturally occurring entities [ISO/IEC/IEEE 15288]

Super Admin – TAI-SDF user who is able to generate checklists, define matrix or create templates

Technical and Organizational Security Measures - means any activities required under Information Security Requirements to access, manage, transfer, process, store, retain, and destroy information or data; to disclose and notify affected parties and to safeguard information or data to ensure availability, integrity, confidentiality, and privacy.

Technical Requirement - the conditions necessary for a system to perform as expected.

Technology Architecture - a description of the structure and interaction of the technology services and technology components. (TOGAF)

Technology Architecture describes the logical software and hardware capabilities that are required to support the deployment of business, data, and application services; this includes IT infrastructure, middleware, networks, communications, processing, standards, etc. (TOGAF)

Telecommunications Network - means a system for carrying communications using guided or unguided electromagnetic energy or both

Third-Party - means Vendor's subcontractors consultants, temporary personnel, contractors, or additional vendors and/or agents acting on behalf of the Vendor. According to the GDPR, third party means a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data [Article 4(10) GDPR].

Threat Assessment — the process of evaluating the actions, events, and behaviors that can cause harm to an asset or organization. Threat assessment is an element of risk assessment and management. (Also known as threat modeling and threat inventory.) [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Threat modeling - a systematic exploration technique to expose any circumstance or event having the potential to cause harm to a system in the form of destruction, disclosure, modification of data, or denial of service. (Source: ISO/IEC/IEEE 24765: 2017)

Threat - any circumstance or event with the potential to adversely impact organizational operations, organizational as-sets, or individuals through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service. [FIPS 200(1)]

TOGAF Architecture Development Method (ADM) provides a tested and repeatable process for developing architectures.

Trust - the reliability and truth of information or the ability and disposition of an entity to act appropriately, within a specified context. [b-ITU-T X.1252]

Trusted third party (TTP) - a security authority or its agent that is trusted (by other entities) concerning some security-relevant activities (in the context of a security policy). [ITU-TX.810]

TRUSTEE - a) is a legal term which, in its broadest sense, is a synonym for anyone in a position of trust and so can refer to any individual who holds property, b) It is also the name of this project and spelled TRUSTEE.

Trustworthy system - a system that meets specific security requirements in addition to meeting other critical requirements

Unauthorized Access — any access or use of a computer system, network or resource which is in violation of the company security policy or when the person or user was not explicitly granted authorization to access or use the resource or system [The A to Z of Cybersecurity Glossary by [Simplicity](#)]

Update - a patch, upgrade, or other modification to code that corrects security and/or functionality problems in software or firmware

Usability –the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.

Use case - specification of a set of actions performed by a system, which yields an observable result that is, typically, of value for one or more actors or other stakeholders of the system. [b-IEC 62559-2]

User – in the eIDAS and SSI framework: the human entity who wants to access a resource controlled by a verifier and obtains credentials from Issuers to this end. In Data Act: means a natural or legal person that owns a connected product or receives a related service or to whom the owner of a connected product has transferred, on the basis of a rental or leasing agreement, temporary rights to use a connected product or receive related services and, where the connected product or related service involves the processing of personal data, the data subject; [Data Act, Article 2(5), version after amendments of the EP, available at: [Texts adopted - Data Act - Tuesday, 14 March 2023 \(europa.eu\)](#)

User agent - the software entity that represents the human User and manages her credentials.

User-centred requirement – a requirement that aims to make the product usable and focused on end user needs and objectives.

User Story – a small story created to achieve a particular objective inside a product.

Verification - confirmation, through the provision of objective evidence, that specified requirements have been fulfilled [ISO 9000:2005]

Virus - a computer program that can copy itself and infect a computer without permission or knowledge of the user.

Visualization - any technique for creating images, diagrams, or animations in order to communicate a message or present information.

Vulnerability - a weakness of an asset or group of assets that can be exploited by one or more threats [ISO 27000]

Whitelist - an approved list of entities provided with a particular privilege or access.

Workshop - an interactive session that involves end-users, researchers, and/or other participants working diligently on a problem or subject.

LIST OF ABBREVIATIONS AND ACRONYMS

ADS	Authenticated Data Structure
AEPD	Agencia Española de Protección de Datos (Spanish Data Protection Authority)
AI	Artificial Intelligence
AIA	Artificial Intelligence Act
AIaaS	Artificial Intelligence Infrastructure as a Service
AIMaaS	AI Models as a Service
AM	Authentication Manager
API	Application Program Interface
AS	Authorization Server
ATR	Accountable Transaction Recorder
B2B	Business-to-Business
B2C	Business-to-Consumer
BDM	Big data management technologies
CA	Certificate Authority
CAP	Consistency, Availability, Partition tolerance
CDSM	Copyright in the Digital Single Market
CDW	Clinical Data Warehouse
CE	Cyber Essentials
CEDS	Common Education Data Standards
CEN	European Committee for Standardization
CJEU	Court of Justice of the European Union
CNIL	Commission Nationale de l'Informatique et des Libertés (French Data Protection Authority)
COVID	Corona Virus Disease
CRF	Case Report Form
CSA	Cloud Security Alliance
CSP	Cloud Service Provider
CSV	Comma-Separated Values
DA	TRUSTEE Dashboard
DbaaS	Database as a service
DDoS	Distributed Denial of Service
DGA	Data Governance Act
DICOM	Digital Imaging and Communications in Medicine
DID	Decentralised Identifier
DL	Data Lake
DLT	Distributed Ledger Technology
DMP	Data Management Plan
DNS	Domain name system
DP	Differential Privacy
DPA	Data Protection Authority
DPIA	Data Protection Impact Assessment
DPM	Data processing and management
DSA	Digital Services Act
DSP	Digital service providers
DSS	Data Security Standard
E2E	End-to-end
EBSI	European Blockchain Services Infrastructure
EC	European Commission
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
edtech	Education Technology Provider
EEA	European Economic Area
EHDS	European Health Data Space

EHR	Electronic Health Record
eIDAS	Electronic Identification, Authentication and Trust Services
EMR	Electronic Medical Record
ENISA	European Union Agency for Network and Information Security
EP	European Parliament
ESA	European Space Agency
ESSIF	European Self-Sovereign Identity Framework
ETP	European Technology Platform
ETSI	European Telecommunications Standards Institute
EU	European Union
FAIR	Findable, accessible, interoperable and reusable
FHE	Fully homomorphic encryption
FL	Federated Learning
GDPR	General Data Protection Regulation
GESS	Global Education Security Standard
GP	General Practitioner
HCI	Human-Computer Interaction
HE	Homomorphic Encryption
HEAT	Homomorphic Encryption Applications and Technology
HEDF	Homomorphic Enabled Data Fusion
HIPAA	Health Insurance Portability and Accountability Act
HL7	Health Level 7
HLEG AI	High-level Expert Group on AI
HLF	Hyperledger Fabric
HLR	High Level requirement
HMI	Human Machine Interface
HPCA	High-Performance Computer Architecture
HRESIA	Human Rights Ethical and Social Impact Assessment
HRIA	Human Rights Impact Assessment
HSS	Homomorphic self-sovereignty
HTTP	Hypertext Transfer Protocol
IaaS	Infrastructure as a Service
ICO	Information Commissioner's Office (UK Data Protection Authority)
ICT	Information and Communication Technologies
ID	Identification
IdP	Identity Provider
IDS	International Data Spaces
IDSA	Internal Data Spaces Association
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IID	Independent and Identically Distributed
IIoT	Industrial Internet of Things
IoT	Internet of Things
IoV	Internet of Vehicles
IP	Internet Protocol
IP (within legal framework)	Intellectual property
IPFS	InterPlanetary File System
IPR	Intellectual property rights
ISMS	Information Security Management System
ISO	International Organization for Standardization
IT	Information Technologies
ITU	International Telecommunication Union
ITU-T	ITU Telecommunication Standardization Sector
JSON	JavaScript Object Notation
JTC	Joint Technical Committee

JWE	JSON Web Encryption
JWT	JSON Web Token
KPI	Key Performance Indicator
KR	Knowledge Repository
KVS	Key-value store
LHE	Levelled Fully Homomorphic Encryption
LINDDUN	Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, and Non-compliance
LOINC	Logical Observation Identifiers Names and Codes
ML	Machine Learning
MVDS	Minimum Viable Database
NASA	National Aeronautics and Space Administration
NCSC	National Cyber Security Centre of the United Kingdom Government
NGO	Non-Governmental Organization
NIS	Network and Information Systems
NIST	National Institute of Standards and Technology
OASIS	Organization for the Advancement of Structured Information Standards
OES	Operator of essential services
OLAP	On-Line Analytical Processing
OneSS	One-Stop-Shop
OWASP	Open Web Application Security Project
P2B	Platform-to-business
P2P	Peer-to-peer
PaaS	Platform as a Service
PbD	Privacy by Design
PCP	Probabilistically checkable proof
PDF	Portable Document Format
PDP	Provable Data Possession
PECKS	Public Key Encryption with Conjunctive Keyword Search
PEKS	Public Key Encryption with Keyword Search
PERKS	Public-key Encryption with Registered Keyword Search
PET	Privacy-enhancing Technology
PHE	Partially Homomorphic Encryption
PHI	Personal Health Information
PIA	Privacy Impact Assessment
PII	Personally Identifiable Information
PKI	Public Key Infrastructure
PoW	Proof of Ownership
PV	Photo-Voltaic
QoS	Quality of Service
R&D	Research and Development
SaaS	Software as a Service
SDLC	Software Development Life Cycle
SDN	Software Defined Network
SHE	Somewhat homomorphic encryption
SIPOC	Suppliers, Inputs, Process, Outputs, and Customers
SL	Swarm Learning
SLA	Service-Level Agreement
SLO	Service-Level Objective
SOAP	Simple Object Access Protocol
SOTA	State-Of-The-Art
SPbD	Security and Privacy by Design
SPoF	Single Point of Failure
SSI	Self-Sovereign Identity/Identification
SSI-HE	SSI Homomorphic Capable Framework
STM	Security and Trust Manager

STRIDE	Spoofing, Tampering, Repudiation, (Information) Disclosure, Denial of Service, and Elevation of Privilege
SWHE	Somewhat Homomorphic Encryption
SWIPO	Switching Cloud Providers and Porting Data
TAI-SDF	Trustworthiness AI Support Design Framework
TDM	Text and data mining
TLD	Top-level-domain
TOM	Technical and Organisational Measure
TRL	Technology Readiness Level
TSP	Trust service providers
TTP	Trusted third party
UIP	Untrusted Infrastructure Provider
UK	United Kingdom
UK GDPR	The United Kingdom General Data Protection Regulation
URL	Uniform Resource Locator
UX	User Experience
VDR	Verifiable Data Registry
VO	Visual Odometry
VPN	Virtual Private Network
W3C	World Wide Web Consortium
WHO	World Health Organisation
WMA	World Medical Association
WP	Work Package
XaaS	Anything as a Service
xAI	Explainable AI
XLS	Microsoft Excel Spreadsheet
XML	Extensible Markup Language
ZKPs	Zero-Knowledge-Proofs

EXECUTIVE SUMMARY

This deliverable describes the achievements of WP2: “Conceptualisation, Use cases and System Architecture” realized during the first year of the project and contains a list of all abbreviations and acronyms and a glossary of terms used in WP2 documents. The content is presented in five parts (high level chapters), addressing all the relevant results based on the Work Package objectives. This includes state of the art analysis, a description of project use cases, a list and description of user-centred, legal, socio-ethical, and technical requirements, and, finally, the initial version of the TRUSTEE system architecture. The work in WP2 started with the definition of the TRUSTEE manifesto, stating the vision, motivation, and principles for TRUSTEE that was especially followed during the development of TRUSTEE architecture.

The mapping of the content to the parts of the D2.1 deliverable is outlined in the section “D2.1 Deliverable Structure” together with the WP2 objectives covered in each of the deliverable parts.

INTRODUCTION

This chapter provides the introduction to the whole document. It presents the mapping of the WP2 objectives to each part of the deliverable (high level chapters), and presents a TRUSTEE Manifesto.

Part I includes the state-of-the-art analysis of various aspects related to the project, such as secure and trusted platforms for data exchange, legal foundation on data and security, homomorphic-capable self-sovereign framework, privacy and security in federated learning, reference architectures, etc..

Part II brings an overview of the TRUSTEE project’s use cases including information about the used datasets (additional to the information provided in D1.6 Data Management Plan available³), data accessibility, security, and protection, as well as an ethical review of the use cases.

Part III contains information on user-centred, legal, and socio-ethical requirements. It describes the methodology behind the extraction of the aforementioned requirement types and User experience (UX) research. This part of the deliverable also brings information about the surveys conducted among TRUSTEE pilot leaders and members, explains TRUSTEE personas and user stories, lists user-centred, legal, and socio-ethical requirements, and provides information on the interconnection between user-centred and legal requirements as well as between legal and socio-ethical requirements.

Part IV describes the technical requirements as well as the process of their identification, including requirements gathering, elicitation, analysis, and specification.

Part V is focused on the architecture of the TRUSTEE system. It describes the process of TRUSTEE architecture development, TRUSTEE architecture principles, and the architecture itself. TRUSTEE Architecture Building Blocks (ABBs) are described and information on the interaction between them is provided.

DELIVERABLE STRUCTURE

This Deliverable D2.1 is of type Report and for clarity and reader’s convenience it is structured in five main parts (or high-level chapters) as presented in the following table where their mapping is given on specific WP2 objectives.

³ Spanakis, E., Sakkalis, V.-F., Forgó, N., Barud, K., Özkayagan-Prändl, H., Sailllant, C. – U., Markakis, E., Papatsaroucha, D. – H., Rosengren, P. – C., Kourtis, A., Xilouris, G. – Z., Hillman, V. – E., Nicolin, A. I. – I., Niskanen, I. – V., Martignano, M.-U., RUTTO, G., TESFAI OGBU, I.-R., Grguric, A. – E., & Lalos, A. – I. (2022). *TRUSTEE Consortium (2022) D1.6 Data management plan available*

Table 0.1 D2.1 parts (high level chapters) mapping to WP2 objectives

D2.1 part	WP2 objectives addressed	Focus
Part I: State of the art	- Perform a systematic gap analysis with a focus on European Data spaces - Indicate the legal framework relevant for TRUSTEE system developers and its users	All collected SOTA, needed for further driving the work of the project
Part II: Use Cases	Define and perform an in-depth analysis of a set of diverse use cases in the field of European data spaces, data security and privacy	TRUSTEE Use Cases
Part III: User Centred, Legal, and Socio-ethical Requirements	- Elicit the TRUSTEE Stakeholders requirements given the outcomes of the analysis performed on User Experience (UX) research outcomes - Indicate legal requirements for TRUSTEE system developers and TRUSTEE users, stemming from the legal acts included in the legal framework presented in Part I. Explore how TRUSTEE solutions could fulfil requirements imposed by current and upcoming European and national legal frameworks regarding European Data spaces - Analyse barriers & contributing factors regarding solutions acceptance & assure TRUSTEE will meet these needs	Personas, User-centred, Legal, and Socio-ethical Requirements, and the interplay between them
Part IV: Technical Requirements	Map User-centred requirements into system requirements and specifications	TRUSTEE technical requirements
Part V: Architecture	- Design and define in detail the overall TRUSTEE architecture of the system. - Identify data flows and interfaces	TRUSTEE architecture

The following figure (Figure 1) shows how the D2.1 parts (high-level chapters) feed each other, starting from State of the art to Use Cases, to User-centred, Legal, and Socioethical requirements, which all lead to the definition of Technical Requirements, which in turn leads to definition of the TRUSTEE Architecture.

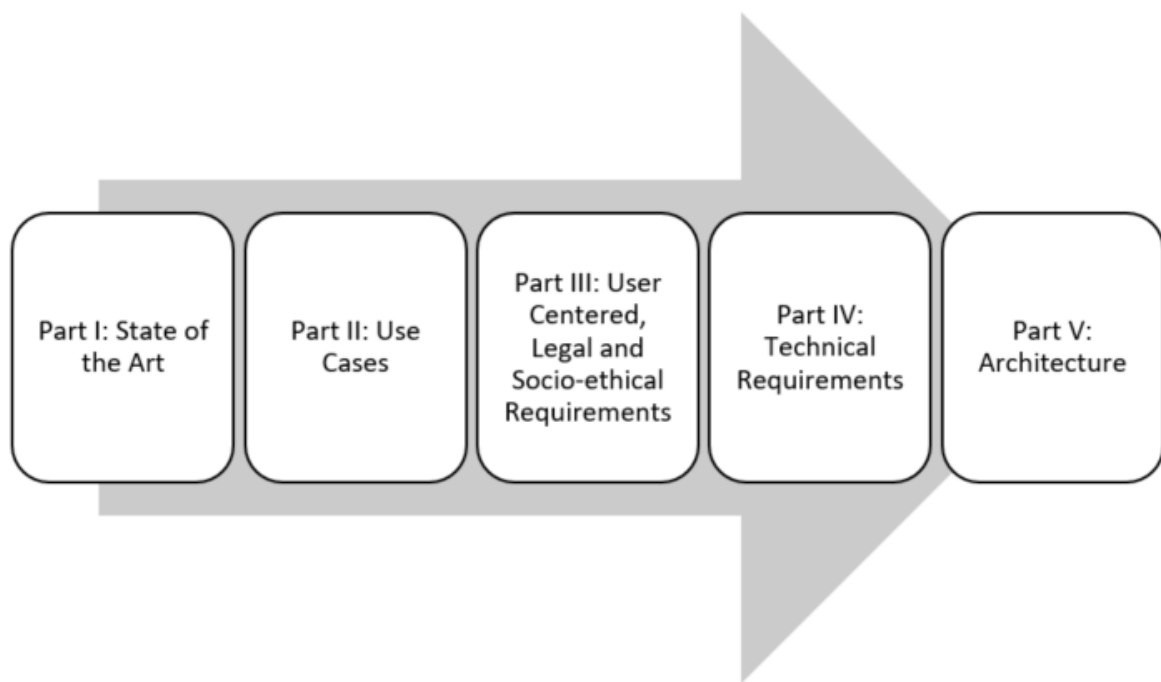


Figure 1: D2.1 parts and how they feed each other

TRUSTEE MANIFESTO

TRUSTEE manifesto was created at the very beginning of the project to make the vision of the project explicit and to further focus the efforts of all TRUSTEE partners into the same direction with common understanding.

The content of the TRUSTEE Manifesto is as follows:

TRUSTEE Manifesto

WHY: This document is prepared by the TRUSTEE project (<https://horizon-trustee.eu/>) consortium to outline its intentions, motivations, and views while setting up to create a Trust & Privacy Preserving Computing Platform for Cross-Border Federation of Data.

WHEN: The project TRUSTEE will run from 1 July 2022 until 31 December 2025.

TO WHOME: The audience of this document is both the consortium participants and other stakeholders including the members of current EU-level initiatives, projects, and organizations working towards making the vision of Open Data Spaces a reality and contributing toward the EU innovation space.

MOTIVATION:

- In the data-driven era, interdisciplinary and dispersed data repositories are inevitable.
- Exposing the data from repositories is non-trivial and requires much effort, especially when dealing with ensuring privacy, security, trust, and accountability.
- Lack of collaborations and data sharing transparency among stakeholders in Europe.

GOALS:

1. TRUSTEE aims to foster collaboration for stakeholders across European data spaces.

2. TRUSTEE aims to realize a trustworthy and privacy-by-design framework for the aggregation of various interdisciplinary data repositories by utilizing state-of-the-art technologies, such as Homomorphic Encryption (HE), and by investigating and leveraging existing EU initiatives and solutions such as GAIA-X and EOSC (European Open Science Cloud).
3. TRUSTEE will facilitate the FAIR (Findable, Accessible, Interoperable, Re-usable) representation of data by providing means to exchange the data by applying HE (enabling computations on its encrypted data without first decrypting them) and other security and privacy-preserving mechanisms.
4. TRUSTEE will employ explainable and trustworthy Artificial Intelligence (AI) processes for efficient and robust use and re-use of data and metadata across interdisciplinary domains
5. TRUSTEE will analyze the ethical, legal, privacy, security, and IPR issues for the collection, storage, processing, querying, analytics, and delivery of data enabling the European single market for data.
6. TRUSTEE will validate the solution for (re)use of various datasets through six different pilots representing different data domains.
7. TRUSTEE will define an agreement management module to ensure legal and GDPR compliance and agreement transparency for accessing the data.

TRUSTEE (architecture) PRINCIPLES:

1. Framework for secure and trustworthy data federation
 - TRUSTEE will be used as an enabler for secure and trustworthy data federation between different Open Data Spaces.
2. Leveraging EU Initiatives and Existing Technologies and standards
 - TRUSTEE will utilize existing and well-established standards, technologies, and initiatives.
3. Data is an asset
 - Data is an asset that has value to the individual or organization and as such should be managed accordingly.
4. Data is secured
 - Data is protected from unauthorized use and disclosure. In TRUSTEE FL is used to prevent undesired information interception and confidentiality loss, and to enable third parties to apply operations on encrypted data without revealing the original data values.
5. Data has a Trustee, Steward, Custodian, and User
 - A Data Trustee has overall responsibility for the data in the organization. Data Steward is responsible for data governance, ensuring regulatory compliance, accuracy, and quality of the (meta)data assets. Data Custodian implements data policies and manages security and access rights controls. The Data User is a consumer of the data, an individual or organization with authorization to access and use the data.
 - TRUSTEE facilitates the secure data exchange between data providers and data consumers.
6. Data should be FAIR
 - Metadata and data should follow FAIR principles (<https://www.go-fair.org/fair-principles/>) to be easy to find, access, and use for both humans and computers.
7. Data should be processed in compliance with legal requirements
 - Data are processed and shared under binding regulations in the field of data protection, data governance, AI application, and online platforms regulations.
8. Applications are technology independent
 - Applications (that will utilize the data accessed via TRUSTEE) are independent of specific technology choices and therefore can operate on a variety of technology platforms.

PART I SUMMARY

The section is dedicated to the survey of the current state of art on all research and technology fields covered by TRUSTEE. A systematic survey of the current technologies, solutions and research in the areas of data management, processing and analytics, security and trust across distributed and federated data sets, privacy, legal and ethical issues of data use and re-use in EU and worldwide, HE and self-sovereign identity management, explainable and trustworthy AI and FAIR and open knowledge extraction, has been conducted to satisfy the WP2 objectives. This survey lays the foundation for TRUSTEE to develop innovative solutions addressing challenges and open questions across different scientific domains (i.e., data science, social, ethical and legal mathematics, computer science and engineering).

The methodology followed during the survey of the current state-of-art and good practices over the different domains based on systematic literature review of related scientific works, tapping into the background knowledge and good practices followed by partners of the TRUSTEE consortium in related areas, as well as exploring technologies and solutions developed by organizations being part of the individual partner's internal network of collaborators. Therefore, the deliverable has achieved an in-depth overview of the current status and shed light into potential shortcomings of the technologies currently used in different areas of the data management and data exchange and AI across the EU.

This deliverable is a live document; hence it will be updated regularly through the lifetime of WP2 in an agile manner. Observations and conclusions drawn during this SOTA survey will be updated, if necessary, in the following months, while extended surveys may be conducted on technologies not considered during this period of the project. Therefore, the development phases will be fuelled by new information regarding SOTA, in a truly agile manner.

The outcomes of the state-of-art survey included in this section are being used as ground truth and baseline technologies for the development of novel solutions across the entire TRUSTEE ecosystem. Regular updates, extensions of the surveyed topics and documentation of newly developed good practices, which will be included in the updated versions of this deliverable, will ensure that TRUSTEE is always aware of the most current and up-to-date technologies.

PART I INTRODUCTION

The European Commission's data strategy, which aims to establish a single market for data, critically depends on the European Data Space. The goal is to maintain control over businesses and individuals who produce the data while ensuring that more data is made available for use in the economy and society. This vision will support data-driven breakthroughs in industries including healthcare, agriculture, energy, and smart cities as well as the EU's competitiveness in the digital economy. The European Data Space is currently being developed, and important components are still being put in place. The construction of shared European data spaces in vital industries, the creation of enabling data governance laws, and the development of a favorable policy climate are a few significant achievements. These steps are intended to create a robust ecosystem that is supported by regulations that are completely consistent with EU principles and laws and allows data to move freely among EU nations, industries, and systems. The main issue is the existence of various data silos and the socioeconomic, legal, and technological restrictions that prevent the free movement of data.

To create the European Data Space, several technological initiatives are now under progress. These include initiatives to improve data infrastructure and platforms, guarantee data quality and dependability, and standardize

data formats and interoperability. With a significant emphasis on privacy-enhancing technology and safe data flow, security is a crucial area of concern. The creation of open, dependable, and trustworthy data platforms and infrastructures, which will guarantee secure data exchange, is a noteworthy endeavour. To guarantee data integrity, confidentiality, and availability, these platforms are supported by technologies like secure cloud storage, edge computing, federated learning, and blockchain.

The TRUSTEE project closely resembles the European Data Space's goals and vision. The project directly contributes to the technological architecture of the data space through the discussion of Secure and Trusted Platforms for data exchange. A further crucial component of the data space is the legal basis on data privacy and security, which guarantees that the project's operation complies with EU legislation. Both the Open and FAIR framework for extracted knowledge and the homomorphic-capable self-sovereign framework are essential to the technological work being done for the data space. These technologies guarantee the security, usability, and accessibility of the data exchanged inside the environment. Furthermore, the development of a safe, interoperable, and user-centric data environment is further aided by the xAI methodologies concentrating on transparency, trustworthiness, and federated Machine Learning (ML). Unlocking the potential of data in Europe is the goal of the innovative initiative known as the European Data Space. This goal is becoming a reality thanks in large part to the efforts of the TRUSTEE project, which places a strong emphasis on data privacy, security, and interoperability.

This section presents comprehensive insights and developments as part of the European Union Horizon project titled "Trustworthy, Reliable and Engaging User Experiences in Digital Europe (TRUSTEE)." It outlines the state-of-the-art innovations in data privacy, security, and technology, striving towards creating a more transparent, secure, and user-focused digital Europe.

- In the first chapter, we delve into the design and development of Secure and Trusted Platforms for data exchange. The focus is on creating a safe ecosystem for data sharing that can significantly enhance the collaborative potential of digital enterprises, underpinning the data-driven innovation across the EU.
- Next, the deliverable outlines the Legal foundation on data privacy and security in the TRUSTEE project. We examine the existing legal frameworks and how they integrate into our project's core. The aim is to ensure that the project's technology infrastructure is aligned with EU laws and regulations, particularly the General Data Protection Regulation (GDPR).
- The third chapter presents our innovative Homomorphic-capable self-sovereign framework. This groundbreaking framework allows for computations on encrypted data, providing a new level of privacy and security for users, thus enhancing trust in digital interactions.
- Subsequently, we introduce the Open and FAIR (Findable, Accessible, Interoperable, Reusable) framework for extracted knowledge. This framework ensures that the insights derived from the TRUSTEE project are readily accessible and usable, promoting knowledge sharing and collaboration.
- We then explore the application of xAI techniques focusing on transparency, trustworthiness, and federated Machine Learning (ML). By enhancing the interpretability and transparency of AI systems, we are aiming to build user trust while leveraging the power of distributed data with federated ML.
- In the chapter on Privacy and Security in Federated Learning: Challenges, Solutions and Applications, we provide a comprehensive analysis of the issues associated with implementing federated learning models. The chapter offers in-depth coverage of the proposed solutions to address these challenges and the potential applications of these models in creating a secure and private data exchange environment.
- We then turn our attention to the State Of The Art on TRUSTEE related Reference Architectures. This chapter provides an overview of cutting-edge technologies and best practices for designing and implementing reliable and secure digital architectures.
- Finally, the TRUSTEE Landscape chapter presents an overview of the project's impact on the broader digital landscape. It explores how TRUSTEE aligns with wider EU objectives for digital innovation, the potential beneficiaries of the project's outcomes, and the future directions for the TRUSTEE project.

THREAT LANDSCAPE

DEFINITION AND PERCEPTION OF RISK, AND REGULATION (GDPR/CLOUD/NIS)

Understanding the constantly evolving threat landscape is crucial for effectively managing the risks associated with data exchange, as organizations must be prepared to mitigate potential threats such as malware, phishing attacks, and advanced persistent threats.

The term “threat landscape” refers to the overall security environment and the various threats that organizations face in their daily operations. In today’s digital age, the threat landscape has become increasingly complex and dynamic, with new threats emerging on a regular basis.

We can enhance our understanding by analyzing statistics from “The Global Risks Report 2023”. In the rapidly evolving landscape of cybersecurity, new technologies and threats surface at a pace never seen before. Alongside this development, there is a noticeable shift in the way businesses and organizations handle cybersecurity. This transformation is primarily influenced by the migration towards cloud-based computing and the ascendancy of digital transformation, encompassing the Internet of Things (IoT) and Artificial Intelligence (AI). However, every transition harbors risks. “The Global Risks Report 2023” identifies cyber-attacks as the top technological risk. For businesses and governments, the primary concerns lie in the sustainability of several areas: the cost-of-living crisis, natural disasters, extreme weather, and the failure to mitigate climate change. It’s crucial to address these areas strategically to mitigate potential vulnerabilities.

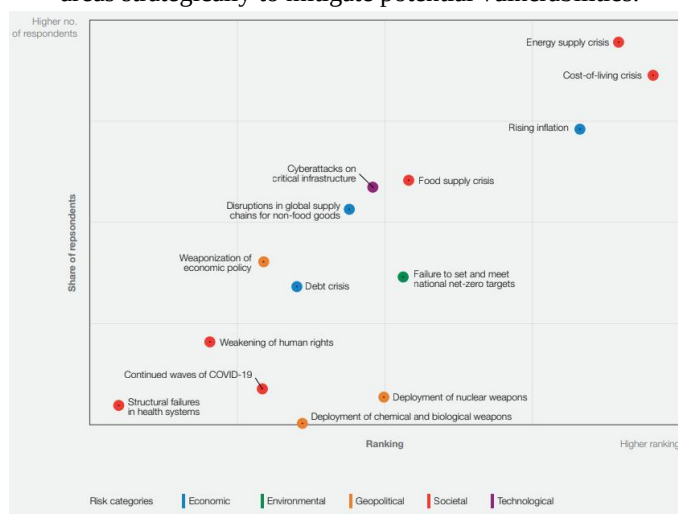


Figure 2 Picture from “The Global Risks Report 2023” – Source: World Economic Forum Global Risks Perception Survey 2022-2023.

A comprehensive understanding of the threat landscape is essential for organizations to effectively protect themselves against cyber-attacks and other security threats. This chapter will explore the various types of threats that organizations face, including:

- **Malware:** Malware refers to malicious software designed to harm or compromise computer systems. This includes viruses, worms, trojans, and other types of malicious software that can infect systems through email attachments, downloads, or other means.

- **Phishing:** Phishing refers to the practice of tricking individuals into revealing sensitive information, such as passwords or credit card details, by posing as a trustworthy entity. This is typically done through email or social media messages that appear to be legitimate but are actually fake.
- **Denial-of-Service (DoS) attacks:** A DoS attack is a type of cyber-attack that seeks to disrupt the normal functioning of a computer system or network by overwhelming it with traffic or other requests.
- **Advanced Persistent Threats (APTs):** APTs are sophisticated and targeted attacks that are designed to gain access to sensitive information over an extended period. These attacks are often carried out by state-sponsored actors or highly skilled cybercriminals.
- The regulations can be considered a threat to cybersecurity, particularly if they are not implemented properly or if they create unintended consequences.

On the one hand, regulations can be an important tool for promoting cybersecurity and protecting sensitive data. Regulations can set minimum standards for security practices, require organizations to disclose data breaches, and impose penalties for noncompliance. This can help ensure that organizations take cybersecurity seriously and invest in the necessary resources to protect their data and systems.

However, there are also risks associated with regulations. For example, regulations can create a false sense of security, leading organizations to believe that compliance with the regulation is sufficient to protect against all cyber threats. Additionally, regulations can be complex and difficult to understand, creating confusion and uncertainty for organizations trying to comply. This can lead to a focus on meeting regulatory requirements rather than implementing best practices that are more effective at protecting against threats.

Furthermore, regulations can be misused by bad actors for malicious purposes. Cybercriminals can use regulations as a pretext to conduct phishing attacks, social engineering scams, and other tactics designed to trick people into revealing sensitive information or providing access to systems.

In the landscape of cybersecurity threats and regulations, the NIS2 Directive holds significant importance. Serving as the EU-wide legislation on cybersecurity, it plays a crucial role in enhancing the overall level of cybersecurity within the European Union. This directive represents an evolution and update of the EU cybersecurity rules that were initially introduced in 2016. The NIS2 Directive, which came into force in 2023, builds upon its predecessor and brings forth new legal measures to address emerging cyber threats and challenges.

One notable aspect of the NIS2 Directive is its applicability scope. It extends its reach to organizations that meet specific criteria. Any organization with more than 50 employees and an annual turnover exceeding €10 million falls within the scope of the directive. Additionally, it also applies to organizations that were previously included in the original NIS Directive. By encompassing a wide range of entities, the NIS2 Directive aims to ensure that cybersecurity measures are adopted across various sectors and industries, regardless of their size or previous inclusion.

Overall, the NIS2 Directive is a significant step forward in bolstering cybersecurity measures within the EU. It represents an updated and comprehensive approach to address the evolving threat landscape, promoting a safer digital environment for organizations and individuals alike. Compliance with the NIS2 Directive is crucial for covered organizations, as it enables them to meet legal obligations, mitigate cybersecurity risks, and contribute to a more secure and resilient digital ecosystem.

Other important regulation that is regulation that arrive from the other side of Oceans, like the Cloud Act: the Clarifying Lawful Overseas Use of Data Act (Cloud Act) is a U.S. law that was enacted in 2018 to facilitate cross-border access by law enforcement to electronic data held by U.S.-based technology companies. The Cloud Act has raised concerns among some privacy advocates, particularly in Europe, due to the potential for conflicts with European Union (EU) data protection regulations.

One potential threat to cybersecurity for Europeans is that the Cloud Act allows U.S. law enforcement to compel U.S.-based technology companies to provide electronic data, including personal data, stored abroad, regardless of the jurisdiction where the data is stored. This means that European data held by U.S. technology companies could be accessed by U.S. law enforcement, potentially without the knowledge or consent of European authorities or individuals. This could lead to violations of European data protection regulations, such as the General Data Protection Regulation (GDPR), which requires that personal data be processed lawfully and with appropriate safeguards.

Additionally, the Cloud Act creates a framework for bilateral agreements between the U.S. and foreign governments, allowing for the sharing of data across borders for law enforcement purposes. This could create opportunities for cybercriminals to exploit weaknesses in the legal framework and gain access to data that would otherwise be protected by local data protection laws.

In closing, the Cloud Act, while perceived as a potential cybersecurity threat by Europeans, is a fairly new regulation, with its full implications yet to be ascertained. Organizations must remain cognizant of the risks it poses and act to safeguard their data in compliance with pertinent data protection legislation.

Regulations represent both a boon and a challenge in the realm of cybersecurity. They serve a crucial role in fostering security and safeguarding sensitive information; however, they can also inadvertently produce counterproductive outcomes or be exploited by cybercriminals. It is thus crucial to adopt regulations judiciously, complementing them with bespoke best practices catering to an organization's unique needs.

The forthcoming section will delve into some burgeoning threats organizations need to recognize, including the Internet of Things (IoT) and ransomware attacks. We will also examine strategies organizations can deploy to curb these threats and bolster their overall security position.

Risk, defined as the potential for harm, loss, or damage resulting from a given activity, decision, or event, is ubiquitous and encountered in diverse contexts such as driving, consuming certain foods, or investing in the stock market.

However, risk perception varies greatly among individuals, influenced by an array of factors including personal experiences, beliefs, and values. Consequently, activities perceived as risky by some may be viewed as innocuous by others.

Risk perception significantly influences our actions and decision-making processes. High-risk perceptions can lead to increased precautionary measures or complete avoidance of the activity, whereas low-risk perceptions might result in engagement without sufficient precautions.

Therefore, understanding risk perception is vital for effective risk management and informed decision-making. This involves considering not only the objective risks associated with a specific activity but also the subjective factors shaping how these risks are perceived.

The Clarifying Lawful Overseas Use of Data (CLOUD) Act is a United States federal law enacted in March 2018 that governs how law enforcement agencies can access electronic data stored by American technology companies outside the United States. It also provides a framework for bilateral agreements between the U.S. and foreign governments regarding cross-border access to electronic data.

Overall, the CLOUD Act is a complex and controversial issue that continues to generate debate and discussion among lawmakers, technology companies, and privacy advocates.

The regulation that should protect the EU citizens against other regulation is the General Data Protection Regulation (GDPR). GDPR is a regulation in the European Union (EU) that aims to protect the privacy and personal data of EU citizens. If an organization processes the personal data of EU citizens, it must comply with the GDPR, regardless

of whether the organization is based in the EU or not. Here are some strategies that organizations can consider protecting data from GDPR:

- Implementing data protection measures: Organizations can implement technical and organizational measures to protect personal data, such as data encryption, access controls, and data minimization. These measures can help prevent unauthorized access, loss, or destruction of personal data.
- Obtaining explicit consent: Organizations must obtain explicit consent from individuals to process their personal data. This means that individuals must be fully informed about how their data will be used and must give their consent freely and without coercion. Organizations should also provide individuals with the ability to withdraw their consent at any time.
- Appointing a Data Protection Officer (DPO): Organizations that process large amounts of personal data or sensitive data are required to appoint a DPO. The DPO is responsible for ensuring that the organization complies with the GDPR and for handling data protection matters.
- Conducting regular audits and assessments: Regularly assessing data protection policies and practices can help organizations identify potential risks and vulnerabilities and ensure compliance with the GDPR.
- Responding to data subject requests: The GDPR gives individuals the right to request access to their personal data, to have their data erased, and to restrict the processing of their data. Organizations must have procedures in place to respond to these requests in a timely and transparent manner.

It's difficult to assign a specific percentage to the threat landscape for data exchange, as the risks and vulnerabilities vary depending on many factors, such as the type of data being exchanged, the security controls in place, and the attackers' motivations and capabilities. However, it's safe to say that the threat landscape is significant and constantly evolving, and organizations must take proactive steps to manage and mitigate the risks. By implementing appropriate security measures, such as encryption, access controls, and regular security assessments, organizations can reduce the likelihood and impact of successful attacks.

NIS 2 VERSUS NIST

In the next chapters we will describe the NIS 2, the NIS Directive, also known as the Directive on Security of Network and Information Systems, is a legislative measure implemented by the European Union (EU) to enhance the cybersecurity and resilience of critical infrastructure and digital services across EU member states. The directive was adopted by the European Parliament and the Council of the European Union on July 6, 2016, and member states were required to transpose it into national law by May 9, 2018.

The primary objective of the NIS Directive is to establish a common level of cybersecurity preparedness and cooperation among EU member states in order to address the increasing cyber threats to essential services and critical infrastructure, such as energy, transportation, healthcare, banking, and digital service providers. By doing so, it aims to ensure the smooth functioning of these services and maintain public confidence in the digital economy.

Key features of the NIS Directive include:

- Scope: The directive applies to operators of essential services (OES), which are identified by member states, as well as to digital service providers (DSPs), including online marketplaces, cloud computing services, and search engines.
- Risk management and incident reporting: OES and DSPs are required to assess the risks they face and take appropriate security measures to prevent and minimize the impact of cybersecurity incidents. They must also report significant incidents to the national competent authorities designated by each member state.
- National competent authorities: Member states are required to designate one or more national competent authorities responsible for overseeing the implementation and enforcement of the directive within their jurisdictions. These authorities play a crucial role in coordinating cooperation and information exchange between member states.

- Cooperation and information sharing: The directive emphasizes cooperation and information sharing among member states and encourages the establishment of Computer Security Incident Response Teams (CSIRTs) to facilitate coordination during cybersecurity incidents.
- Penalties and enforcement: Member states must establish penalties for non-compliance with the directive and ensure that effective enforcement mechanisms are in place.

The NIS Directive aims to create a harmonized framework for cybersecurity across the EU, promoting a coordinated approach to prevent, detect, respond to, and recover from cyber incidents. It seeks to strengthen the overall resilience of critical infrastructure and digital services, fostering trust and ensuring a more secure digital environment within the EU.

ATTACK METHODOLOGIES FOR DATA EXFILTRATION

To consider in the threat landscape the risk and potential attack that refer the data exfiltration: data exfiltration refers to unauthorized or malicious activities aimed at stealing or extracting sensitive information from a target organization or system. There are various attack methodologies that adversaries may employ to accomplish data exfiltration. Here are some common techniques:

- Network-Based Attacks:
 - Man-in-the-Middle (MitM) Attacks: Intercepting network communications to capture sensitive data.
 - Packet Sniffing: Capturing and analyzing network traffic to extract data packets containing valuable information.
 - Network Eavesdropping: Monitoring network transmissions to gather sensitive data.
- Malware-Based Attacks:
 - Keyloggers: Installing malicious software that records keystrokes to obtain passwords, credentials, or other sensitive information.
 - Remote Access Trojans (RATs): Deploying malware that provides unauthorized remote access to the target system, allowing the attacker to retrieve data.
 - Backdoors: Creating secret entry points in a system to enable unauthorized access and data exfiltration.
- Insider Threats:
 - Malicious Insiders: Employees or trusted individuals within an organization who intentionally steal and exfiltrate sensitive data.
 - Unintentional Insiders: Employees who accidentally expose or transfer sensitive data without malicious intent, often due to lack of awareness or proper security practices.
- Phishing and Social Engineering:
 - Phishing Attacks: Deceptive emails, messages, or websites that trick users into disclosing sensitive information, such as login credentials.
 - Spear Phishing: Highly targeted phishing attacks that focus on specific individuals or organizations to gain access to sensitive data.
- Physical Attacks:
 - Unauthorized Access: Physically breaching the premises to gain access to physical storage devices or systems containing sensitive data.
 - Tampering: Manipulating or compromising physical infrastructure, such as network devices or storage systems, to facilitate data exfiltration.
- Cloud-Based Attacks:
 - Cloud Service Exploitation: Exploiting vulnerabilities in cloud platforms or misconfigurations to gain unauthorized access and exfiltrate data stored in the cloud.
 - API Exploitation: Abusing insecure APIs to extract sensitive data from cloud-based applications or services.

This is not exhaustive, but just a few examples of attack methodologies used for data exfiltration.

SPIFFEE AND SPIRE AUTHENTICATION

In the context of discussing the threat landscape, it is crucial to emphasize the significance of authentication as an essential tool in maintaining a secure environment. Authentication serves as a fundamental component of any robust cybersecurity strategy, enabling organizations to verify and validate the identities of users, devices, and services accessing their systems. Within the realm of authentication, workload authentication plays a particularly relevant role.

Workload authentication refers to the process of verifying the identity of individual workloads or software components within a distributed system. As organizations increasingly adopt cloud-native architectures and microservices-based applications, ensuring the authentication of these workloads becomes paramount. Workload authentication helps prevent unauthorized access, reduces the risk of compromised services, and safeguards critical assets from potential cyber threats.

In this regard, one solution that we suggest for workload authentication is the use of:

- Secure Production Identity Framework for Everyone (SPIFFE)
- SPIFFE Runtime Environment (SPIRE)

SPIFFE and SPIRE are open-source projects that provide a standardized approach to workload authentication in distributed systems.

One defines a set of specifications and standards for workload identity, while the other implements those specifications in a practical runtime environment. Together, they enable organizations to establish a secure and trusted identity framework for their workloads, irrespective of the underlying infrastructure or platform.

SPIFFE and SPIRE are two different open-source projects related to secure communication and identity management in distributed systems.

SPIFFE is an open standard that provides a way to authenticate and authorize communication between services in distributed systems. It defines a set of standards for issuing, verifying, and managing identities in distributed systems, so that services can communicate securely and reliably without relying on IP addresses or other unreliable identifiers.

SPIRE is an implementation of the SPIFFE standard that provides a set of tools and services for managing SPIFFE identities and enforcing authentication and authorization policies in distributed systems. It includes a set of agents that run on each host in the system and coordinate with a central server to issue and manage identities for services running on that host.

Both SPIFFE and SPIRE are designed to provide a secure, scalable, and flexible way to manage identities and secure communication in distributed systems, which is becoming increasingly important as more organizations adopt cloud-native architectures and microservices. SPIFFE and SPIRE are primarily designed to authenticate and authorize communication between services in distributed systems, rather than authenticating individual users.

By adopting SPIFFE and SPIRE, organizations can achieve several benefits. Firstly, it offers a uniform and scalable approach to workload authentication across heterogeneous environments, making it easier to manage and maintain a consistent security posture. Additionally, Spiffe and Spire provide strong identity guarantees, ensuring that workloads can securely communicate with each other and that only authorized entities can access sensitive resources.

PROCEDURES, POLICIES, TECHNICAL GUIDELINES, TOOLS AND RECOMMENDATIONS FOR DATA PRIVACY AND CONFIDENTIALITY

In this chapter we will develop the Procedures, Policies, Technical Guidelines, Tools and Recommendations for data privacy and confidentiality starting from the Hewlett Packard Enterprise (HPE) model that leverages on NIST and other standards. As an example, HPE ESRM is a security reference model proposed by HPE. It is a holistic approach to security that aims to protect the entire IT environment, from the edge (network perimeter) to the cloud.

The HPE ESRM reference model is based on four fundamental pillars:

- **Edge Protection:** This pillar focuses on security at the edge, where network connections and peripheral devices meet the outside world. It includes the physical and logical security of endpoints, protection of local area networks (LAN) and wireless networks, as well as access and identity management.
- **Entry Point Protection:** This pillar focuses on the security of entry points such as firewalls, security gateways, and data security devices. It ensures that incoming data is protected from external threats and that system access is controlled and secure.
- **Data Protection:** This pillar addresses data security throughout its lifecycle, from creation to storage and transmission. It includes data encryption, cryptographic key management, access control, backup security, and monitoring of data access activities.
- **Cloud Protection:** This pillar focuses on the security of cloud services and resources used by an organization. It includes protecting access to cloud services, API (Application Programming Interface) security, data segregation, compliance with regulations, and managing threats in the cloud.

The HPE ESRM reference model provides a framework for understanding and systematically addressing IT security challenges along the entire path from the edge to the cloud. It offers guidelines and best practices for designing, implementing, and managing robust and integrated security solutions that protect an organization's resources, data, and applications. Here are procedures and policies that organizations can implement to ensure secure data exchange and data privacy:

- **Develop a data classification scheme:** Classify data based on its level of sensitivity and ensure that appropriate security controls are implemented based on the classification. For example, highly sensitive data may require encryption or additional access controls.
- **Implement encryption:** Encrypt data in transit and at rest to protect it from unauthorized access. Use strong encryption algorithms and ensure that encryption keys are managed securely.
- **Use secure communication channels:** Use secure communication channels such as SSL/TLS or VPN to ensure that data is transmitted securely between parties.
- **Implement access controls:** Implement access controls to ensure that only authorized individuals can access sensitive data. Use strong authentication mechanisms such as two-factor authentication or biometric authentication to verify user identity.
- **Establish data exchange agreements:** Establish data exchange agreements with third-party organizations to define the terms and conditions for data exchange, including security requirements and responsibilities.
- **Conduct regular security assessments:** Conduct regular security assessments to identify and address vulnerabilities and risks in the data exchange process. Use third-party security assessments or penetration testing to identify potential weaknesses in security controls.
- **Develop incident response procedures:** Develop incident response procedures to address security incidents or breaches. Establish a clear chain of command and ensure that incident response plans are regularly tested and updated.
- **Employees awareness:** Train employees on the importance of secure data exchange and the procedures and policies in place. Conduct regular security awareness training to ensure that employees understand the risks and their role in protecting sensitive data.

PROCEDURES AND POLICIES FOR SECURE DATA EXCHANGE (C.I.A.)

The key is to implement a comprehensive security framework that considers the risks associated with data exchange and the needs of the organization, of course for any model we will the famous triad CIA, the triad ensure that data is protected from unauthorized access, tampering, and disruption.

The C.I.A. (Confidentiality, Integrity, Availability) triad is a widely used security framework that can be applied to secure data exchange. Here are some procedures and policies that utilize the C.I.A. triad for secure data exchange:

- **Confidentiality:** Protect data from unauthorized access and disclosure. Use encryption to protect data in transit and at rest. Implement access controls to ensure that only authorized individuals can access sensitive data. Use strong authentication mechanisms to verify user identity. Develop data classification schemes to identify and protect highly sensitive data.
- **Integrity:** Ensure that data is accurate and has not been tampered with. Use digital signatures or other methods of data authentication to ensure data integrity. Implement data backup and recovery procedures to ensure that data can be restored to its original state in the event of data corruption or loss. Implement version control to track changes to data and ensure that only authorized changes are made.
- **Availability:** Ensure that data is available when needed. Implement redundancy and failover mechanisms to ensure that data remains available in the event of system failure or other disruptions. Develop disaster recovery plans to ensure that data can be quickly restored in the event of a disaster or other disruptions. Conduct regular system maintenance and monitoring to identify and address potential performance or availability issues.

TLS/DATA ENCRYPTION/AUTHENTICATION/AUTHORIZATION

To achieve the triad, we need to introduce the data in motion protection (TLS), data encryption, authentication, and authorization contribute to the C.I.A. triad as follows:

- **Confidentiality:** TLS and data encryption provide confidentiality by protecting data from unauthorized access or disclosure.
- **Integrity:** TLS, authentication, and authorization mechanisms ensure the integrity of data by preventing unauthorized tampering or modifications.
- **Availability:** By securely authenticating and authorizing users, these mechanisms help maintain the availability of data and resources by preventing unauthorized access that could lead to disruptions or denial of service.
- **TLS (Transport Layer Security):** TLS is a cryptographic protocol that provides secure communication over networks, such as the internet. It ensures confidentiality and integrity by encrypting data during transmission and verifying the authenticity of the communicating parties. TLS prevents unauthorized interception or tampering of data, safeguarding its confidentiality, and maintaining its integrity.
- **Data Encryption:** Data encryption involves converting plain, readable data into ciphertext using cryptographic algorithms. Encryption ensures confidentiality by making the data unreadable to unauthorized individuals or entities. Encrypted data can only be decrypted using the appropriate decryption key, providing protection against data breaches or unauthorized access.
- **Authentication:** Authentication verifies the identity of users, systems, or entities attempting to access a resource or service. It ensures that only authorized individuals or systems are granted access. Authentication mechanisms typically involve validating credentials such as usernames, passwords, digital certificates, or biometric factors. By confirming identities, authentication helps prevent unauthorized access and ensures that data is accessed by legitimate entities only.
- **Authorization:** Authorization determines the level of access or privileges granted to authenticated users or systems. It involves defining access control policies and permissions based on user roles, groups, or other criteria. Authorization ensures that users can only access the resources or perform actions they are authorized to, protecting against unauthorized activities, and preserving data integrity.

TECHNICAL GUIDES: DESCRIPTION OF ENCRYPTION, AUTHENTICATION, AND AUTHORIZATION MECHANISMS (SPIFF FOCUS)

Encryption mechanisms are cryptographic techniques used to protect sensitive information by converting it into an unreadable form, known as ciphertext. Encryption provides confidentiality and helps safeguard data from unauthorized access or disclosure. Here are some common encryption mechanisms:

- **Symmetric Encryption:** Symmetric encryption uses a single shared secret key to both encrypt and decrypt data. The same key is used for both operations, making it faster than other encryption methods. However, the challenge lies in securely distributing and managing the shared key among the communicating parties.
- **Asymmetric Encryption (Public Key Encryption):** Asymmetric encryption uses a pair of keys - a public key and a private key. The public key is freely distributed, while the private key is kept secret. Data encrypted with the public key can only be decrypted with the corresponding private key. Asymmetric encryption provides a secure method for exchanging encryption keys and enables digital signatures.
- **Hash Functions:** Hash functions are not encryption mechanisms per se but are often used in conjunction with encryption. They are commonly used to verify data integrity, password storage, and digital signatures.
- **Hybrid Encryption:** Hybrid encryption combines both symmetric and asymmetric encryption.
- **Homomorphic Encryption:** HE allows computations to be performed on encrypted data without the need for decryption. It enables data to remain encrypted while still being processed, thus preserving privacy and confidentiality.

As previously describe the encryption is the countermeasure of Cloud Act and other extra UE regulation, As the CLOUD Act is a U.S. federal law, it applies to U.S. technology companies that store electronic data outside of the United States. As such, protecting data from the CLOUD Act may be challenging, but there are some strategies that individuals and organizations can consider:

- **Encrypting data:** By encrypting data before it is stored in the cloud, it becomes more difficult for law enforcement agencies to access the data without the appropriate decryption keys. However, it is important to note that the CLOUD Act may require technology companies to provide access to decrypted data under certain conditions.
- **Using non-U.S. technology companies:** Using technology companies that are not subject to the CLOUD Act, and do not have a presence in the United States, may help to protect data from U.S. law enforcement agencies. However, it is important to carefully evaluate the privacy and security policies of any technology company before entrusting them with sensitive data.
- **Storing data locally:** Storing data on local servers or other devices, rather than in the cloud, may provide greater control over access to the data. However, this approach may be less convenient than cloud storage and may require additional resources to maintain.

Ultimately, protecting data from the CLOUD Act requires careful consideration of the risks and benefits of different approaches, and ongoing evaluation as the legal and technological landscape evolves.

TOOLS AND RECOMMENDATIONS: PKI AND SPIFF HSM, KMS, DISK/STORAGE ENCRYPTION TOOLS

There are essential tools used in modern cybersecurity to ensure secure key management, protect sensitive data, and establish trusted communication channels:

- PKI (Public Key Infrastructure),
- HSM (Hardware Security Module),
- KMS (Key Management System)
- encryption data at rest with KMIP (Key Management Interoperability Protocol).

- Authentication/Authorization (SPIFFE and SPIRE)

It's also a suggestion to have the previous tools in the Trustee project.

SPIFFE and SPIRE can help enforce security in a workload and prevent data leaks by providing a secure way to authenticate and authorize communication between services in a distributed system

However, it is possible to use SPIFFE and SPIRE to authenticate users in certain scenarios, such as when the users are themselves running services in a distributed system. In this case, each user service could be issued a SPIFFE identity, which would be used to authenticate and authorize communication between the user service and other services in the system.

In addition, SPIFFE and SPIRE could be used as part of a broader identity and access management (IAM) system to authenticate and authorize users, especially in the context of a distributed application running in a cloud-native environment. For example, a centralized IAM system could issue SPIFFE identities to user-facing services, which could then be used to enforce authentication and authorization policies across the distributed application. However, this would require additional layers of authentication and authorization to be built on top of SPIFFE and SPIRE.

By using SPIFFE and SPIRE to issue and manage identities for services, it becomes possible to ensure that only authorized services can communicate with each other, and that communication is encrypted and authenticated. This can help prevent unauthorized access to sensitive data, as well as protect against man-in-the-middle attacks and other security threats.

In addition, SPIFFE and SPIRE can be used to enforce fine-grained access controls based on the identities of services and their associated permissions. This can help ensure that only authorized services have access to specific data or resources and can help prevent accidental or intentional data leaks.

However, it's important to note that while SPIFFE and SPIRE can provide a strong foundation for security in a distributed system, they are not a complete solution in themselves. Other security measures such as encryption, firewalls, and access controls may also be necessary depending on the specific requirements of the workload and the environment in which it is deployed.

TRUSTWORTHINESS-BY-DESIGN FRAMEWORKS FOR MONITORING, AUDITING AND CONFORMITY CHECK

AI systems are playing a growing role as part of decision and control systems in economic and social sectors. Given the superior performance of AI technologies as compared to traditional IT technologies, technical, academic and business communities can benefit from AI applications by reducing labor and promoting social good. However, recent research indicate that AI can cause unintentional harm to humans by making unreliable decisions in safety-critical scenarios or undermining fairness by inadvertently discriminating against a group or groups⁴. Consequently, a large body of literature on trustworthy AI has emerged; trustworthy AI can be defined as programs and systems built to solve problems like a human, which bring benefits and convenience to people with no threat or risk of harm. A systematic approach is needed to build new AI applications in a trustworthy manner; this approach must take into consideration the end-to-end data science and AI lifecycle. The AI lifecycle consists of a sequence of stages that bring together various personas, such as Data Scientist, AI Engineer, Business owner and Product Manager. The stages of the lifecycle shall include⁵:

⁴ Trustworthy AI: A Computational Perspective, Haochen Liu and Yiqi Wang, Michigan State University, USA

⁵ Trustworthy AI: From Principles to Practices, Bo Li, Peng Qi, Bo Liu, Shuai Di, Jingen Liu, Jiquan Pei, Jinfeng Yi, and Bowen Zhou 2022 Association for Computing Machinery

- Scope and plan: this stage guides the prioritization of the use case scenarios and development of an AI action plan leading to business requirements definition.
- Data Management: during this stage data are collected and pre-processed, for instance features selection.
- Algorithm Design and Development: during this stage data science teams explore and prepare the data, build, train and test their AI/ML models. Bias should be checked before building the model. Model robustness, explainability and other aspects can be tested. Once these tests complete successfully, an MLOps pipeline allows for the model and related assets to be moved from the development environment to a pre-production validation environment.
- Verification, Validation and Deployment: this stage involves validation of the model and deployment into production. Validation activities could be performed by a team other than the one that built the model, such as the organization's model validation or model risk management team or an external entity. This team validates quality, robustness, fairness, etc. and generates validation reports.
- Operation: in this stage AI/ML models are monitored. Quality, robustness, and fairness are periodically monitored according to the business need.

Considering that AI Trustworthiness needs to be constantly monitored along all AI stages, a new industrial AI workflow has been introduced. Recent studies extended the DevOps paradigm to MLOps to cover the ML lifecycle of data preparation, training, validation and deployment.

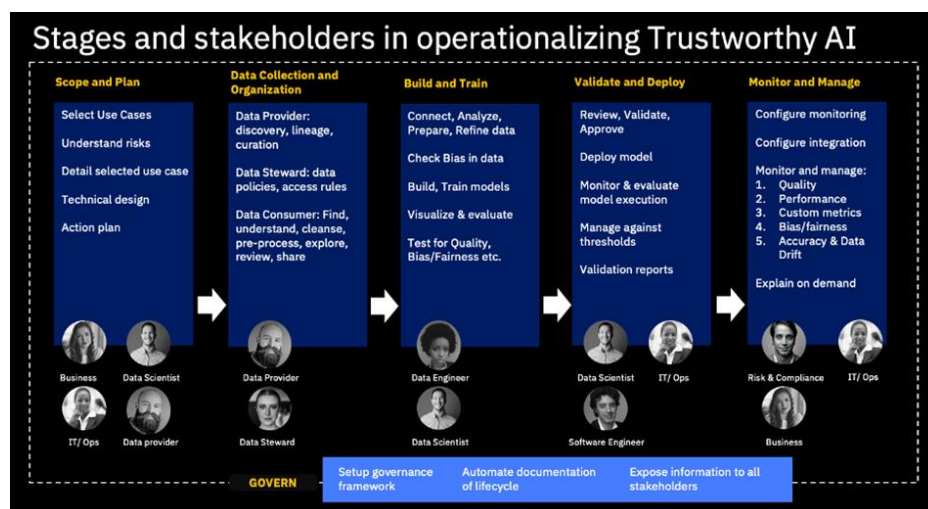


Figure 3: AI system lifecycle vs actors per each stage⁶

Considering that the majority of AI applications target safety and security relevant sectors, a sufficient level of robustness, security and auditability needs to be achieved and corresponding technical guidelines and standards need to be developed. Despite this lack of standards, an emerging set of standards have been issued and depicted in the following picture.

⁶ Foundations of trustworthy AI: Operationalizing trustworthy AI <https://www.ibm.com/blog/operationalizing-trustworthy-ai/>

Topic	Document
Reliability & Robustness	› ISO/IEC NP 24029: Assessment of robustness of neural networks [102]
	› ITU-T F.AI-DLFE & F.AI-DLPB: deep learning software framework evaluation methodology & metrics and evaluation methods for DNN processor benchmark [103]
	› ETSI DTR INT 008 (TR 103 821): AI in test systems and testing AI models, definitions of quality metrics [104]
	› DIN SPEC 92001-2: AI life cycle processes and quality requirements, Part2 robustness [105]
	› ITU Focus Group on "Artificial Intelligence for Health", e.g. [106] and [107]
Safety	› ISO/CD TR 22100-5: safety of machinery, Part 5: implications of embedded AI – ML [108]
	› ISO 26262: road vehicles – functional safety ([18], see also IEC 61508-1:2010 [17], ISO 21448 [109])
	› IEEE P2802: standard for performance and safety evaluation of AI medical devices – terminology [110]
	› ISO/IEC AWI TR 5469: AI – functional safety and AI systems [111]
Security	› ISO/SAE 21434: road vehicles – cybersecurity engineering ([112], s.a. ISO/CD 24089 [113], ISO/IEC 23894 [114])
	› ETSI ISG SAI: Several documents covering problem statement, threat ontology, security testing and mitigation strategies for general AI systems, e.g. [115]
	› NISTIR 8269: a taxonomy and terminology of adversarial ML [116]

Figure 4: Current standards, guidelines, and procedures to ensure AI systems auditability

In TRUSTEE, a Trustworthiness Support Design Framework (TAI-SDF) will be developed: this framework is based on an auditing workflow that aims to support AI researchers and practitioners to verify if AI systems/products are compliant with AI quality requirements stemming from latest European Union standards.

Recent research studies examined several Trustworthiness components; however TAI-SDF only considers the following ones:

- Compliance: assessment of the conformity of a process/product/service respect to the relevant standards/policies/regulations
- Fairness: the data and algorithms of a system prohibit bias and prevent unfair treatment towards any segment of the population
- Explainability: when technical requirements call for human explanation, system actions and decisions are justified in understandable ways.
- Robustness: statistically speaking, a system deals well with outliers and environmental changes. It recognizes changes it cannot handle and shuts down gracefully rather than crashing, spewing nonsense or going completely away.
- Security: a system cannot be hacked or hijacked and resists attacks.
- Privacy: a system does not violate user or subject privacy

EXISTING TOOLKITS IMPLEMENTING TRUSTWORTHY AI

According to the latest ethics guidelines for AI provided by the European Union⁷, a trustworthy AI system should meet four ethical principles: respect for human autonomy, prevention of harm, fairness and explicability. Beyond ethics guidelines issued by European Commission, in May 2019 OECD launched the adoption of the OECD AI Principles where Trustworthy AI refers to AI systems embody these principles, that is AI systems that respect human rights and privacy transparent, explainable, robust, secure and safe; and the actors involved in their development and use remain accountable.⁸ The OECD AI Principles provide five values-based principles for the responsible stewardship of trustworthy AI:

- Inclusive growth, sustainable development and wellbeing: stakeholders should engage in creating credible AI that can contribute to inducing outcomes that are beneficial for people, as well as for the planet.

⁷ Nathalie A Smuha. 2019. The eu approach to ethics guidelines for trustworthy artificial intelligence. Computer Law Review International 20, 4 (2019), 97–106.

⁸ TOOLS FOR TRUSTWORTHY AI: A FRAMEWORK TO COMPARE IMPLEMENTATION TOOLS, OECD DIGITAL ECONOMY PAPERS June 2021, No.312

- Human-centred values and fairness: the values of human rights, democracy, and rule of law should be incorporated throughout the AI system’s lifecycle, while allowing human intervention through safeguard mechanisms.
- Transparency and explainability: AI actors that develop or operate AI systems should provide information to foster an overall understanding of the systems among stakeholders, in which people affected by AI systems could comprehend the outcome and challenge the decision when needed.
- Robustness, security and safety: AI systems need to function appropriately while ensuring traceability, while AI actors need to apply systematic risk management approaches to mitigate safety risks.
- Accountability: AI actors should respect the principles and should be accountable for the proper operation of AI systems.

In June 2020, the working group OECD Network of Experts on AI conducted a survey and classified tools, understood as instruments and structured methods that can be leveraged by others to facilitate their implementation of the AI Principles (e.g. toolkits to check for biases or robustness in an AI system, risk management guidelines, educational material).⁹ Within the plethora of tools, the most relevant ones include technical tools that aim to address specific AI-related issues for each Trustworthiness dimension, namely fairness, transparency, explainability and robustness.

Objective	Tool	Description
Fairness	AT&T software System to Integrate Fairness Transparently (SIFT)	Software system to integrate mechanised and human-in-the-loop components in bias detection, mitigation, and documentation of projects at various stages of the machine learning lifecycle.
	Microsoft Fairlearn	Open-source toolkit to assess and improve the fairness of machine learning models. Contains an interactive visualisation dashboard and bias mitigation algorithms to help navigate trade-offs between fairness and model performance.
	LinkedIn Fairness Toolkit (LiFT)	Open-source toolkit to enable measurement of fairness according to a multitude of fairness definitions in large-scale machine learning workflows.
	Google What-If Tool	Open-source software tool to visually inspect and explore machine learning model performance and data across multiple hypothetical situations, with minimal coding required.
	IBM AI Fairness 360	Open-source toolkit to help detect and mitigate unwanted bias in machine learning models and datasets. Provides approximately 70 metrics to test for biases, and 10 algorithms to mitigate bias in datasets and models.
Transparency	IEEE Standard for Transparency of Autonomous Systems	Technical standard to describe measurable and testable levels of transparency, so that autonomous systems can be assessed and levels of compliance determined.
	Google Model Card Toolkit	Documentation framework for sharing the essential facts of a machine learning model in a structured, accessible way, providing an overview of what the model is intended to do, how it was architected, trained, and its limitations.
Explainability	Google Cloud Explainable AI service	Software to help developers get explanations on the outcomes of their models. Can be applied to the AI models trained on tabular, image, and text data. Not open source.
	IBM AI Explainability 360 Toolkit	Open-source toolkit of algorithms, code, guides, tutorials, and demos to support the interpretability and explainability of machine learning models.
	Microsoft InterpretML	Open-source toolkit containing machine learning interpretability algorithms to help understand model predictions.
Robustness	IBM Adversarial Robustness 360 Toolkit	Open-source toolkit for machine learning security. It provides tools to evaluate, defend, certify and verify machine learning models and applications against the adversarial threats of evasion, poisoning, extraction, and inference.

Figure 5: Technical tools from Working Group implementing Trustworthy AI

It is worth to notice that the aforementioned technical tools are vertically focused on each Trustworthiness component; conversely, TAI-SDF will be a proprietary methodologically framework enabling auditing,

⁹ “Tools for trustworthy AI: A framework to compare implementation tools”, OECD digital economy papers, no. 312, June 2021

continuously monitoring of AI/ML models and data in compliance with latest AI regulations issued by European Commission and ensuring AI models are fair, robust, reliable and secure.

SECURITY AND TRUST MANAGERS

This section highlights the existing approaches and tools for creating a secure and trusted consent managers for facilitating the data sharing. In this study, we classify the existing approaches into four classes:

- Consent management framework: It highlights the existing works with their use cases describing the consent management framework.
- Ontology-based consents: It describes the existing ontology for creating and managing a consent.
- Blockchain-based consents: It describes the how the blockchain is used for recording consents
- Knowledge graph: It shows the knowledge graph model for consents.

We describe in detail these classes.

CONSENT MANAGEMENT FRAMEWORK

An extension of the access control language and algorithm to implement consent and data requests has been described in the literature¹⁰. The extension is based on the Open Digital Rights Language policy language, which allows expressing rich rules, and the Data Privacy Vocabulary, which permits invoking privacy and data protection-specific terms. The user expresses his/her data policy by yes/no questions. They used Solid for the data access control. Solid, the emerging technology for organizing data in decentralized stores, relies on a simple authorization mechanism for granting access to data. Solid's personal online datastores (Pods) are ideal for keeping personal data, as they allow individuals to represent the access permissions in a very simple manner using Access Control Language (ACL) expressions. However, the user cannot to express more complex policies beyond a yes/no. In addition, the actions are not aligned to GDPR concepts, and the data is a data subjects data and not a sample of datasets.

Furthermore, several use cases have been presented¹¹ of permissions from the All of Us informed consent document and translate these permissions into semantic web rule languages expressions to extend and operationalize informed consent ontology. However, the rules and the ontological translations are only examples, used to show proof of concept, and in their infancy. The rules are not generalized to accurately express document-level permission entities from informed consent.

Additionally, empirical experiments have been performed¹² with two major consent management providers in the EU: Quantcast and OneTrust and paired with a legal analysis. They discussed the requirements for consent management providers to be qualified as processors and as controllers and concluded that such status has to be assessed with regard to each specific data processing activity. However, there is no use-case to apply their approach.

¹⁰ Beatriz Esteves, Harshvardhan J. Pandit, Víctor Rodríguez-Doncel: ODRL Profile for Expressing Consent through Granular Access Control Policies in Solid. EuroS&P Workshops 2021: 298-306

¹¹ Muhammad Amith, Marcelline R. Harris, Cooper Stansbury, Kathleen Ford, Frank J. Manion, Cui Tao: Expressing and Executing Informed Consent Permissions Using SWRL: The All of Us Use Case. CoRR abs/2108.10221 (2021)

¹² Cristiana Santos, Midas Nouwens, Michael Toth, Nataliia Bielova, Vincent Roca: Consent Management Platforms Under the GDPR: Processors and/or Controllers? APF 2021: 47-69

Moreover, a consent request user interface prototype has been introduced¹³, which is based on the GDPR requirements and the interpretation of those requirements. Their prototype provides transparency regarding personal data processing, more control via a customization, and, based on the results of our usability evaluation, improves user comprehension with respect to what data subjects actually consent to. However, the approach only focused on health data.

Finally, the MAGIC project at the University Medicine Greifswald has been presented¹⁴. It was to design and implement a structured exchange format for modular consent templates. The format had to take common standards in the area of digital informed consent management into account. The proposed approach is to use a consent of the patient for the medical research. However, the contribution remains in its infancy.

ONTOLOGY-BASED CONSENTS

An ontology has been developed¹⁵ based on hierarchical vocabulary of human and machine-readable data use terms that consistently and unambiguously represents a dataset's allowable data uses. Using the ontology in data management and access facilitates researchers' discovery and access of relevant datasets. Ontology annotations increase the FAIRness of datasets and support data linkages using common data use profiles when integrating the data for secondary analyses. However, it is not clear how to use an ontology for facilitating the discovery of, request for, and access to datasets? How to use terms and definitions for consented dataset? The approach only focused on Biomedical and sensitive human data.

Moreover, an ontology has been proposed¹⁶ for facilitating the development of data sharing and use agreements. The ontology contains four top level classes, 20 sub-classes, 33 sub-categories and 17 simple properties for categories applicable at various levels. The ontology is used for data sharing across organizations. However, the approach is not GDPR compliance, no tool and no use-case are proposed.

Finally, GConsent has been proposed¹⁷, an OWL2-DL ontology for representing information associated with consent for the GDPR. GConsent uses PROV-O and its GDPR-specific extension GDPRov to model provenance of consent, and GDPRtEXT to link concepts to the relevant text within the GDPR. The ontology, its documentation, and this paper is available at <https://w3id.org/GConsent> under CC-by-4.0 license. However, the consent is modelled as an artefact, which has states (given, not given, refused, withdrawn).

BLOCKCHAIN-BASED CONSENTS

¹³ Olha Drozd, Sabrina Kirrane: Privacy CURE: Consent Comprehension Made Easy. SEC 2020: 124-139

¹⁴ Martin Bialke et al., MAGIC: once upon a time in consent management—a FHIR ® tale. Bialke et al. J Transl Med (2018) 16:256, <https://doi.org/10.1186/s12967-018-1631-3>

¹⁵ Jonathan Lawson, Moran N. Cabili, Giselle Kerry, ..., Jaime M. Guidry Auvil, Tommi H. Nyro"nen, Me´lanie Courtot. The Data Use Ontology to streamline responsible access to human biomedical datasets. Cell Genomics 1, 100028 November 10, 2021

¹⁶ M. Greene, Sam Grabus, Jane Greenberg:

DARSI: An Ontology for Facilitating the Development of Data Sharing and Use Agreements. CoRR abs/2109.11046 (2021)

¹⁷ Harshvardhan J. Pandit, Christophe Debruyne, Declan O'Sullivan, and Dave Lewis. GConsent - A Consent Ontology based on the GDPR. ADAPT Centre. <https://openscience.adaptcentre.ie/ontologies/GConsent/docs/ontology>

In current literature, a discussion¹⁸ can be found for the progress and opportunities of remote patient monitoring using blockchain technologies and its two primary frameworks: Ethereum and Hyperledger Fabric. In fact, they provided a thorough overview of blockchain technology and its two widely frameworks, (i) Ethereum and (ii) Hyperledger Fabric. They proposed a system architecture for data management of remote patient monitoring using Ethereum as a model. However, the proposed approach is in its infancy and not mature. The consent management tool is not developed and only focused on a user data related to patient which is a remote patient monitoring.

Moreover, an e-consent model (electronic consent instead of a consent using paper forms) has been introduced¹⁹, which uses a purpose-based access control scheme. It is implemented by a blockchain system using Hyperledger Fabric. All metadata of patient records, consents, and data access are written immutably on the blockchain and shared among participant organizations. The blockchain provides reliability and availability with transparency and traceability. Patients can create, update, and withdraw their consents in the blockchain. However, the consent model is dedicated for consent management both for patients and healthcare organizations. The blockchain recodes the consent as well as the patient records which is not adapted when a gigabyte of data is shared among organizations which the case of our work.

Moreover, a framework has been proposed²⁰ that covers emerging need regarding the privacy of users in the IoT ecosystem. It aims to provide an environment to help users retain control over their personal data in the IoT ecosystem, in line with the GDPR requirements. It controls consents regarding access to their personal data in the IoT ecosystem and exercise their rights defined by GDPR. A blockchain infrastructure is used to ensure the integrity of personal data processing consent. However, there is no monitoring of consents. The work only focused on the use of consent in a blockchain. The use of GDPR compliance is not clear.

Additionally, PrivySharing has been introduced²¹, namely, a blockchain-based secure and privacy-preserving data sharing mechanism for smart cities. The proposed approach ensures that personal user data is kept confidential, securely processed and is exposed to the stakeholders need-to-know basis as per user-defined rules embedded in smart contracts. However, the blockchain is used for Privacy-Preserving and Secure Data Sharing in Smart Cities and not well adapted to consent management. However, the blockchain is used for Privacy-Preserving and Secure Data Sharing in Smart Cities and not well adapted to consent management.

Finally, smart contract for consent-driven and double-blind data sharing on the Hyperledger Fabric blockchain platform has been proposed²². However, there are trade-offs between preserving privacy and promoting transparency.

KNOWLEDGE GRAPH

¹⁸ Md Jobair Hossain Faruk, Hossain Shahriar, Maria Valero, Sweta Sneha, Sheikh Iqbal Ahamed, Mohammad Ashiqur Rahman: Towards Blockchain-Based Secure Data Management for Remote Patient Monitoring. ICDH 2021: 299-308

¹⁹ Dara Tith, Joong-Sun Lee, Hiroyuki Suzuki, W. M. A. B. Wijesundara, Naoko Taira, Takashi Obi, Nagaaki Ohshima. Patient Consent Management by a Purpose-Based Consent Model for Electronic Health Record Based on Blockchain Technology. Healthcare Informatics Research 2020;26(4):265-273.

²⁰ Konstantinos Rantos, George Drosatos, Konstantinos Demertzis, Christos Ilioudis, Alexandros Papanikolaou, Antonios Kritsas: ADvoCATE: A Consent Management Platform for Personal Data Processing in the IoT Using Blockchain Technology. SecITC 2018: 300-313

²¹ Imran Makhdoom, Ian Zhou, Mehran Abolhasan, Justin Lipman, Wei Ni: PrivySharing: A blockchain-based framework for privacy-preserving and secure data sharing in smart cities. Comput. Secur. 88 (2020)

²² Kumar Bhaskaran, Peter Ilfrich, Dain Liffman, Christian Vecchiola, Praveen Jayachandran, Apurva Kumar, Fabian Lim, Karthik Nandakumar, Zhengquan Qin, Venkatraman Ramakrishna, Ernie G. S. Teo, Chun Hui Suen: Double-Blind Consent-Driven Data Sharing on Blockchain. IC2E 2018: 385-391

A knowledge graph-based tool for GDPR consent compliance verification in Business-to-Business (B2B) and Business-to-Consumer (B2C) data sharing has been recently proposed and developed²³. It binds GDPR's legal basis to data sharing contracts. They proposed an ontology and knowledge graph for consents that can be reused in various cases and domains. The proposed approach only focused on two use-cases: connected car and smart city domains. However, there is no approach to translate GDPR legal basis into ontology and knowledge graph.

Finally, an approach has been proposed²⁴ to visualise consent decisions (i.e., what happens to an individual's data after consent is given) in the domain of vehicle sensor data sharing. The visualisation portrays certain aspects of the knowledge graph, mainly how data is processed after consent is given. It aims to raise individuals' awareness of what it means to consent and the implications that follow. It helps users to inspect data sharing activities at any time after consent was given. However, the proposed approach only focused on vehicle data owners sharing.

SUMMARY

All the existing approaches are consent oriented which means that the consent is established between a data subject and an organization (data controller) which differs from Trustee perspectives. In fact, Trustee plans to share the data between organizations. Thus, the legal term and concept of "consent" will not be used in this case. However, the legal term of "agreement" is suitable and used for the rest of the document. Moreover, to the best of our knowledge, there is no work related to agreement management among organizations. Even is consent management tools few works are mature which provides end-to-end tool in a real use-case. There is no tool which covers the GDPR compliance, secure and privacy requirements.

SEMANTIC ONTOLOGIES (ACC)

Semantic ontologies are an important part of the work within TRUSTEE, as "ensuring semantic interoperability (enabler of machine computable logic, inferencing, knowledge discovery and data federation between information systems)"²⁵ is one of the main objectives of the TRUSTEE project in the context of TRUSTEE. Additionally, the 3rd main objective of the project is "Novel Big Data management and analytics infrastructure to facilitate use and re-use of data in data spaces across interdisciplinary science fields and business sectors"²⁶ by:

- Semantic interlinking, harmonization tools and quality checking mechanisms specifications available.
- Development of Algorithms for the semantic annotation of data.
- Algorithms for data, text, visual analysis of big data volumes integrated into the final solution.

Finally, two of TRUSTEE's main technological components (i.e., Data Fusion Middleware and Data Lake) will consist of and handle semantic annotation, interoperability and querying semantic data. More specifically, the Data Fusion's Ontologies and Semantic Representation layer will be responsible for the transformation of the collected data into knowledge, by applying common/standardized semantic ontologies on the data, correlated with the appropriate contextual information²⁷.

²³ Amar Tauqeer, Anelia Kurteva, Tek Raj Chhetri, Albin Ahmeti, Anna Fensel: Automated GDPR Contract Compliance Verification Using Knowledge Graphs. Inf. 13(10): 447 (2022)

²⁴ Christof Bless, Lukas Dötlinger, Michael Kaltschmid, Markus Reiter, Anelia Kurteva, Antonio J. Roa-Valverde, Anna Fensel: Raising Awareness of Data Sharing Consent Through Knowledge Graph Visualisation. SEMANTiCS 2021: 44-57

²⁵ TRUSTEE Grant Agreement Part A, p. 11.

²⁶ TRUSTEE Grant Agreement Part B, p. 5.

²⁷ TRUSTEE Grant Agreement Part B, p. 13.

SEMANTIC ONTOLOGIES AND SEMANTIC ANNOTATION

In the era of big data, the Semantic Web has become an essential framework for managing and sharing data in a machine-understandable format. Semantic Web, i.e., Web 3.0., allows data exchange and integration by providing the tools for machines to understand the meaning of data; such tools include ontologies and vocabularies that assign reasoning mechanisms and formal representation of the knowledge behind the data²⁸.

Semantic ontologies have been developed to enable computers to interpret and understand the meaning of data, which is crucial for efficient data integration, interoperability, and sharing²⁹. These ontologies are comprised of sets of classes, properties and relationships, and are used to annotate different domains with specific meaning. Semantic annotation is the process of adding explicit and formal semantic metadata to web resources, such as documents, images, videos, etc., with the aim of enabling their automatic processing and understanding by machines. The semantic annotations are typically expressed using a formal language, such as RDF, OWL, or RDFS^{30,31}, and adhere to a specific ontology or vocabulary that provides a shared understanding of the concepts and relationships between them. The annotations can be added manually, or automatically by machine learning algorithms that use natural language processing techniques to extract the relevant concepts and relationships from unstructured data³².

The annotation of these domains' data with concepts from semantic ontologies enables easier representation, faster processing and targeted querying that generates intelligence knowledge. This also allows interoperability of heterogeneous data across different domains, an essential asset for data sharing and reusability²⁹.

SEMANTIC ONTOLOGIES FOR TRUSTEE DOMAINS

This section presents the review carried out on semantic ontologies that represent classes, properties and relationships for each of the 5 main domains within TRUSTEE, i.e., Health, Education, Space, Energy and Automotive. More specifically, the review identified 57 semantic ontologies in total that include classes, properties and relationships for all 5 domains, as well as domain-independent ontologies and ontologies relevant to the Security and Trust Manager component, and more specifically the Agreement Manager sub-component within T4.6. Each list of ontologies is presented in the following Tables, where the most relevant for the TRUSTEE concepts findings were included, based on the datasets to be provided by each pilot.

The study also includes a list of domain-independent semantic ontologies, that can be used for any datasets, as they represent generic information (e.g., person, organization, location, event, documents etc.) (see

²⁸ Chen, B., Zhao, Y., Yang, J., & Wang, F. (2021). Ontology engineering for the Semantic Web: a systematic review. *Journal of Intelligent Information Systems*, 56(2), 207-236. This paper provides a comprehensive review of ontology engineering for the Semantic Web, highlighting its importance for enabling data integration, interoperability, and knowledge sharing.

²⁹ Anwar, A. A. (2022). A survey of semantic web (Web 3.0), its applications, challenges, future and its relation with Internet of things (IoT). In *Web Intelligence* (No. Preprint, pp. 1-30). IOS Press.

³⁰ Gupta, R., & Malik, S. K. (2022, December). Visualizing Semantic Web Data using Various Tools Focusing RDF, OWL and SPARQL. In *2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART)* (pp. 1456-1460). IEEE.

³¹ Ding, L., Kolari, P., Ding, Z., & Avancha, S. (2007). Using ontologies in the semantic web: A survey. *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*, 79-113.

³² Dimitrov, D., Baran, E., Fafalios, P., Yu, R., Zhu, X., Zloch, M., & Dietze, S. (2020, October). Tweetscov19-a knowledge base of semantically annotated tweets about the covid-19 pandemic. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (pp. 2991-2998).

Table 0.7). These ontologies can either be used to fill in gaps from other ontologies when a relevant class cannot be found, or to amplify the semantic description of a concept that has already been annotated with a domain-specific ontology.

Table 0.2 Semantic ontologies for the Health domain

Semantic Acronym	Ontology	Semantic Ontology Title	Semantic Ontology Description
HL7 ³³		Health Level Seven (HL7) Reference Information Model	healthcare information (patients, encounters, and other aspects of healthcare delivery)
CIDO ³⁴	Coronavirus Infectious Disease Ontology		knowledge related to the COVID-19 pandemic and other coronavirus infectious diseases, such as clinical symptoms, laboratory tests, epidemiology, and public health interventions
SNOMED CT ³⁵		Systematized Nomenclature of Medicine - Clinical Terms	clinical findings, procedures etc.
FHIR ³⁶	Fast Healthcare Interoperability Resources		patients, encounters, and other clinical concepts
TMO ³⁷		Translational Medicine Ontology	diseases, phenotypes, biological processes, molecular functions, cellular components, and clinical interventions.
CTO ³⁸		Core Ontology of Clinical Trials	clinical trials data, drugs, devices, procedures, outputs, demographic characteristics, medical history
DOID ³⁹		Human Disease Ontology	diseases and disorders, symptoms, causes, risk factors, treatments, therapies

³³ <https://bioportal.bioontology.org/ontologies/HL7>

³⁴ <https://bioportal.bioontology.org/ontologies/CIDO>

³⁵ <https://bioportal.bioontology.org/ontologies/SNOMEDCT>

³⁶ <http://hl7.org/fhir/R4/>

³⁷ <https://bioportal.bioontology.org/ontologies/TMO>

³⁸ <https://bioportal.bioontology.org/ontologies/CTO>

³⁹ <https://bioportal.bioontology.org/ontologies/DOID>

OMRSE ⁴⁰	Ontology of Medically Related Social Entities	socioeconomic status, race, ethnicity, gender, education, patients relationships with doctors, health delivery systems etc.
CTV3 ⁴¹	Read Codes, Clinical Terms Version 3	medical procedures, medicine, medical procedures, patient registration, appointment scheduling etc.
CPT ⁴²	Current Procedural Terminology	medical, surgical, and diagnostic procedures
ICD-10 ⁴³	International Statistical Classification of Diseases and Related Health Problems	diseases and related health problems
meSH ⁴⁴	Medical Subject Headings	indexing and searching biomedical literature and categorizing health-related concepts
OpenEHR ⁴⁵	OpenEHR Reference Model	patients, encounters, and other aspects of healthcare
OGMS ⁴⁶	Ontology for General Medical Science	biomedical data, anatomical structures, diseases, treatments
SHR ⁴⁷	Student Health Record Ontology	history of disease, family history of disease, public examinations

Table 0.3 Semantic ontologies for the Education domain

Semantic Acronym	Ontology	Semantic Ontology Title	Semantic Ontology Description
ESCO ⁴⁸		European Skills, Competences, qualifications and Occupations	skills, competences, qualifications, and occupations

⁴⁰ <https://bioportal.bioontology.org/ontologies/OMRSE>

⁴¹ <https://bioportal.bioontology.org/ontologies/RCD?p=classes&conceptid=X78FG>

⁴² <https://bioportal.bioontology.org/ontologies/CPT>

⁴³ <https://bioportal.bioontology.org/ontologies/ICD10>

⁴⁴ <https://bioportal.bioontology.org/ontologies/MESH>

⁴⁵ <https://ckm.openehr.org/ckm/archetypes/1013.1.120>

⁴⁶ <https://bioportal.bioontology.org/ontologies/OGMS>

⁴⁷ <https://bioportal.bioontology.org/ontologies/SHR>

⁴⁸ <https://esco.ec.europa.eu/en>

IEEE LOM ⁴⁹	Learning Object Metadata	characteristics of learning resources, instructional materials, videos, assessments
SIF ⁵⁰	School Interoperability Framework	information from student information systems and learning management systems
VIVO ⁵¹	VIVO Core Ontology	students, academics, research organizations, publications
TEACH ⁵²	Teaching Core Vocabulary Specification	Courses, materials, teachers, students, buildings
LRMI ⁵³	Learning Resource Metadata Initiative	educational resources
CEDS ⁵⁴	Common Education Data Standards	student demographics, course information, and assessment results

Table 0.4 Semantic ontologies for the Space domain

Semantic Acronym	Semantic Ontology	Semantic Ontology Title	Semantic Ontology Description
QuakeML ⁵⁵	QuakeML ontology		seismic events such as their location, magnitude, depth, and focal mechanism
EONET	Event Ontology for Earthquakes and Tsunamis		earthquake and tsunami events, including their location, magnitude, time, and impact, fault lines and tectonic plates
SWEET ⁵⁶	Semantic Web for Earth and Environmental Terminology		Earth and environmental science concepts Land form, impact crater, volcanic crater, erosional crater, crater lake

⁴⁹ <https://lov.linkeddata.es/dataset/lov/vocabs/lom>

⁵⁰ <http://specification.sifassociation.org/Implementation/2.3/index.html>

⁵¹ http://akswnc7.informatik.uni-leipzig.de/dstreitmatter/archivo/vivoweb.org/ontology--core/2020.06.10-212153/ontology--core_type=generatedDocu.html

⁵² <http://linkedscience.org/teach/ns#>

⁵³ <https://www.dublincore.org/specifications/lrmi/1.1/>

⁵⁴ <https://ceds.ed.gov/whatIsCEDS.aspx>

⁵⁵ <https://quake.ethz.ch/quakeml/>

⁵⁶ <https://github.com/ESIPFed/sweet/blob/master/README.md>

LCD	Lunar Crater Database ontology	types of lunar craters (complex craters, simple craters, and impact basins, their size, depth, and location and relationships such as the "is parent of" relationship between an impact basin and the craters located within it
GeoNames ⁵⁷	GeoNames Ontology	geographical features, craters, place names, locations
SEO	Solar Event Ontology	types of solar events, including solar flares, e.g. "solar flare," "flare class," "flare duration," and "flare peak time"
SWO	Space Weather Ontology	types of space weather phenomena, including solar flares, e.g. "solar flare," "flare intensity," and "flare duration"
SDM	Solar Data Model	solar events, e.g. "flare event," "flare peak," and "flare onset" and properties such as "flare intensity" and "flare duration."
GeoSciML ⁵⁸	Geological Survey Markup Language	geological features (e.g. craters, rocks, materials, location)

Table 0.5 Semantic ontologies for the Automotive domain

Semantic Acronym	Ontology	Semantic Ontology Title	Semantic Ontology Description
AutoOnt ⁵⁹		Automotive Ontology	vehicles, such as make, model, and year, as well as the components and systems within a vehicle
Vehicle ⁶⁰		Vehicle Ontology	vehicles information (e.g. brand, capacity, acceleration etc.)

⁵⁷ <http://www.geonames.org/ontology/documentation.html>

⁵⁸ <http://geosciml.org/>

⁵⁹ <http://www.automotive-ontology.org/>

⁶⁰ <https://enterpriseintegrationlab.github.io/icity/Vehicle/doc/index-en.html>

VISS ⁶¹	Vehicle Information Specification	Service	vehicle information, signals from sensors on control units within a vehicle's network
AIAG	Automotive Industry Action Group		attributes of a vehicle and its components
VSS ⁶²	Vehicle Signal Specification		the state of a vehicle
VSO ⁶³	Vehicle Sales Ontology		cars, boats, bikes, and other vehicles

Table 0.6 Semantic ontologies for the Energy domain

Semantic Acronym	Ontology	Semantic Ontology Title	Semantic Ontology Description
ESO		Energy System Ontology	energy systems and their components, including energy generation, transmission, distribution, storage, and consumption.
SEAS		Photovoltaic ontology	PV installations and PV systems
KM4C			main elements of a photovoltaic system
REEGLE		Renewable Energy and Energy Efficiency	renewable energy and energy efficiency technologies, including PV installations (components, properties, and performance of PV systems, as well as their environmental and economic impacts)
PSO		Photovoltaic System Ontology	technical aspects of PV systems, including their components, configurations, and performance
ODP		Energy Information Administration's (EIA) Open Data Platform Vocabulary	energy-related data.
SGO		Smart Grid Ontology	components and operations of a smart grid

⁶¹

[https://w3c.github.io/automotive/spec/VISSv2_Transport.html#:~:text=The%20Vehicle%20Information%20Service%20Specification,Vehicle%20Signal%20Specification%20\(VSS\).](https://w3c.github.io/automotive/spec/VISSv2_Transport.html#:~:text=The%20Vehicle%20Information%20Service%20Specification,Vehicle%20Signal%20Specification%20(VSS).)

⁶² https://covesa.github.io/vehicle_signal_specification/introduction/overview/

⁶³ <http://www.heppnetz.de/ontologies/vso/ns>

OEM	Ontology for Energy Management	energy management systems and their components
ERO	Energy Resources Ontology	energy resources, including their generation and consumption

Table 0.7 Domain-independent semantic ontologies

Semantic Acronym	Ontology	Semantic Ontology Title	Semantic Ontology Description
IAO		Information Artifact Ontology	data sets, software tools, experimental protocols, research publications
FOAF ⁶⁴		friend of a friend	information about people, organizations and their relationships
Schema.org ⁶⁵		Schema.org	various types of data, e.g. person, organization, event, location
DCMI ⁶⁶		Dublin Core Metadata Initiative	resources and educational materials
SKOS ⁶⁷		Simple Knowledge Organization System	concepts, collections, members, notes

Semantic Acronym	Ontology	Semantic Ontology Title	Semantic Ontology Description
ODRL ⁶⁸		Open Digital Rights Language	permission, conditions and obligations for any digital content and service

⁶⁴ <http://xmlns.com/foaf/0.1/>

⁶⁵ <https://schema.org/>

⁶⁶ <https://www.dublincore.org/resources/glossary/ontology/>

⁶⁷ <https://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html>

⁶⁸ <https://www.w3.org/TR/odrl-model/>

W3C PROV-O ⁶⁹	Provenance Ontology	provenance information for tracking the origin, ownership, and usage of data
DPV ⁷⁰	Data Privacy Vocabulary	privacy policies, consent, and personal data processing activities, as well as other privacy-related concepts
GConsent ⁷¹	Gconsent Ontology	consent-related data between different applications and systems based on GDPR
GDPRtEXT ⁷²	General Data Protection Regulation (GDPR) Ontology	personal data, datasets, data processing, legal basis, controller agreement, consent
P-Plan ⁷³	Provenance Planning Ontology	as plans, steps, inputs/outputs

SEMANTIC INTEROPERABILITY FRAMEWORKS

EU has set up frameworks and principles that encourage semantic annotation of data to improve data interoperability and reuse.

The European Open Science Cloud (EOSC)⁷⁴ is a framework that aims to provide open access to research data and promote the use of semantic technologies, such as RDF ontologies, to enhance data discovery, sharing, and reuse. This initiative is led by the European Union (EU), with the objective to create a virtual space that brings together existing scientific resources and data from across Europe, and make scientific data easily accessible and reusable, fostering collaboration and innovation across different disciplines and countries.

The EOSC framework is built on a set of key principles, including openness, interoperability, and reusability. The idea is to provide a single point of access to a wide range of data and services, from raw data and software to tools for data analysis and visualization.

Another important aspect of the EOSC is the development of common standards and protocols for data management and sharing. This includes the use of common data models and metadata standards, as well as the development of common data access protocols. This will allow researchers to easily discover, access, and reuse data from different sources, regardless of where it is stored or how it is formatted.

⁶⁹ <https://www.w3.org/TR/prov-o/>

⁷⁰ <https://w3c.github.io/dpv/dpv/>

⁷¹ <https://openscience.adaptcentre.ie/ontologies/GConsent/docs/ontology>

⁷² <https://openscience.adaptcentre.ie/ontologies/GDPRtEXT/deliverables/docs/ontology#ShouldBeCommonlyUsedFormat>

⁷³ <http://vocab.linkeddata.es/p-plan/index.html>

⁷⁴ <https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1>

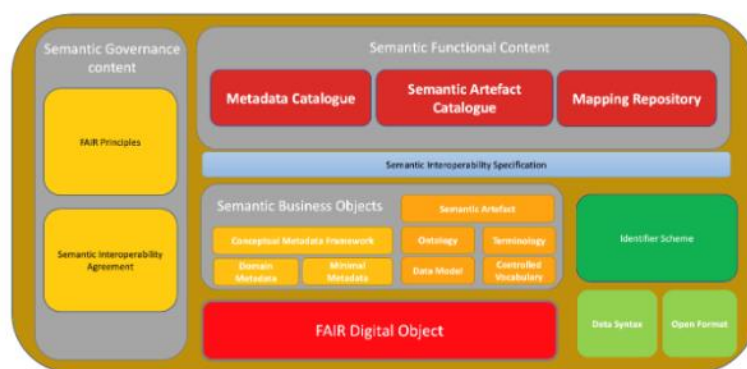


Figure 6: Semantic view of the EOSC framework

The FAIR principles (Findability, Accessibility, Interoperability, and Reusability) are a set of guidelines that aim to make data findable, accessible, interoperable, and reusable. Semantic annotation of data can help achieve these principles by adding machine-readable metadata that can facilitate data discovery, integration, and analysis.

The European Interoperability Framework (EIF) is a framework that provides guidance on how to achieve interoperability in the European public sector. Semantic annotation of data can help achieve this goal by enabling better integration and sharing of data across different public services and agencies.

LEGAL FOUNDATION ON DATA PRIVACY AND SECURITY IN TRUSTEE

This section of the Deliverable presents the initial analysis of the legal framework for the TRUSTEE project in the context of ensuring privacy and security standards, being the outcome of the work conducted under T1.5 – Legal and Data Protection Constraints in TRUSTEE and T2.2 - Legal Framework for data manipulation and AI processing. These legal foundations indicate the relevant specific regulations and are the basis for the definition of further legal requirements for TRUSTEE, discussed in detail in Part II of the Deliverable. They aim to guide the TRUSTEE system developers and its users (Data providers, Model providers, Consumers and Developers) throughout the lifecycle of the development and maintenance of the system, as well as indicate what requirements they would need to fulfil when the system is available outside of the research environment to achieve legal and ethical compliance with the binding regulations in the sphere of, in particular, data protection, cybersecurity, eIDAS and AI application, data governance and digital services.

It must be emphasised that the legal framework presented below is an initial framework, and it may be subject to modifications, which will be further included in D2.2 and D2.3.

FUNDAMENTAL RIGHTS AND PRIMARY LAW

The fundamental rights envisaged under the Charter of Fundamental Rights of the EU lie at the heart of the EU *acquis* as its guiding principles. We see a clear reference to the Charter in this manner, for instance, in the GDPR⁷⁵, Data Governance Act⁷⁶ and the proposal for the AI Act⁷⁷. The Charter was adopted in 2000 and has been a legally

⁷⁵ GDPR Recital 1 and 4.

⁷⁶ Data Governance Act, Recital 46.

⁷⁷ AI Act Proposal, Recital 28.

binding instrument of the EU primary law since 2009 with the entry into force of the Treaty of Lisbon. With this, the Charter binds the EU member states as well as the EU institutions in their acts within the scope of EU law.^{78 79}

For the purposes of TRUSTEE, the two main fundamental rights enshrined under the Charter are the right to respect the private and family life of a person (Article 7) and the right to protection of personal data concerning a person (Article 8). The Charter dictates the following:

“Article 7 - Respect for private and family life: Everyone has the right to respect for his or her private and family life, home and communications.

Article 8 - Protection of personal data:

1. Everyone has the right to the protection of personal data concerning him or her.
2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
3. Compliance with these rules shall be subject to control by an independent authority.”⁸⁰

It is expected that the parties processing personal data would balance their legitimate interest with the fundamental rights and freedoms of the data subjects and evaluate if the latter takes precedence over their own legitimate interests. This sort of balancing act is also borne out of the fact that the fundamental rights enshrined under Article 7 and 8 are not absolute rights but they “must be considered in relation to their function in society and be weighed against other fundamental rights”, with due regard to the principle of proportionality.⁸¹ Thus, any interference with these rights are prohibited except for the limitations allowed for under Article 52(1) of the Charter.

Next to the Charter, this document refers to three main instruments within the framework of human rights law:

-*Universal Declaration of Human Rights (1948)* - the Declaration stands as a list of guarantees for the protection of human rights and acts as a foundation of main principles of equality, justice and freedom for all individuals.

⁷⁸ Grainne De Burca, ‘After the EU Charter of Fundamental Rights: The Court of Justice as a Human Rights Adjudicator?’, *Maastricht Journal of European and Comparative Law*, 2013, Vol 20(2), p 169.

⁷⁹ When mentioning the binding effect of the Charter on EU member states and the EU institutions, we mostly refer to the vertical effect of fundamental rights, which concerns the relations between a state and its individuals. This scope is also stipulated under Article 51(1) of the Charter. Thus, in principle, this fundamental rights framework does not apply horizontally, meaning to relations between private parties, but can still have an indirect horizontal effect. See Ciacchi, Aurelia Colombi, ‘European Fundamental Rights, Private Law, and Judicial Governance’, in Hans Micklitz (ed.), *Constitutionalization of European Private Law: XXII/2*, Collected Courses of the Academy of European Law (Oxford, 2014; online edn, Oxford Academic, 22 May 2014) and Prechal, Sacha. “Horizontal Direct Effect of the Charter of Fundamental Rights of the EU.” *Revista de Derecho Comunitario Europeo*, vol. 2020, no. 66, Centro Estudios Políticos Constitucionales, 2020, pp. 407–26, <https://doi.org/10.18042/cepc/rdce.66.04>.

⁸⁰ Charter of Fundamental Rights of the EU, Articles 7 and 8.

⁸¹ C-439/19, *Latvijas Republikas Saeima (Penalty points)*, 22 June 2021, EU:C:2021:504, para. 105 and C-311/18, *Facebook Ireland and Schrems*, 16 July 2020, EU:C:2020:559, paras 172-176.

-*OHCHR International Covenant on Civil and Political Rights* - It expands on the rights and freedoms in the Universal Declaration of Human Rights and form together with it the “International Bill of Human Rights”.⁸²

-*The Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights)* - it is the binding instrument for some of the rights and freedoms that were drawn from the Universal Declaration of Human Rights.

Article 12 of the Universal Declaration of Human Rights, similar to Article 7 of the Charter of the Fundamental Rights of the EU and Article 8 of the ECHR, protects individuals against any arbitrary interference with their privacy, family, home or correspondence and grants protection of the law against any such interference. There is no specific mention of protection of personal data of individuals in these instruments. However, this right can be derived from Article 12 of the Declaration, from its reference to the right to privacy. Article 17 of the International Covenant on Civil and Political Rights also recognizes the right to privacy in the same vein, similar to the ECHR with the right to respect for private and family life.⁸³

Article 29(2) of the Declaration is almost akin to Article 52 of the Charter of the Fundamental Rights of the EU. It reminds the necessity of not to limit the exercise of the mentioned rights arbitrarily but only by law for certain reasons, such as to allow proportional and balanced exercise of the rights and freedoms of others.

For TRUSTEE, though there may be no need to directly address these instruments and articles as they are inferred from the text and spirit of the *EU acquis* in its entirety, the protections envisaged under these articles should always act as guiding principles for the project and in the development the TRUSTEE solution by TRUSTEE system developers, as well as its use by different types of users, especially before any processing of personal data.

DATA PROTECTION LEGAL FRAMEWORK – SECONDARY LEGISLATION

The focus of Section 1.2 is put on the General Data Protection Regulation, being the basis for assessing and explaining data protection approach which should be embedded in TRUSTEE solution as a standard one.

INTRODUCTION

Personal Data, Data Related To Health, Data Concerning Minors

⁸² <https://www.ohchr.org/en/treaty-bodies/ccpr/background-international-covenant-civil-and-political-rights-and-optional-protocols>

⁸³ See also De Gregorio, G. (2022). Digital constitutionalism across the Atlantic. *Global Constitutionalism*, 11(2), 297-324, p. 305: “Prior to the Charter, the protection of freedom of expression, privacy and personal data in the European context was based not only on the domestic level but also on the European Convention on Human Rights (Convention). The Strasbourg court has played a crucial role in extending the protection of fundamental rights online as well as underlining the constitutional challenges coming from digital technologies. For instance, the court has underlined the risks for freedom of expression online. Likewise, in the field of data, it has underlined the relevance of protecting personal data evolving from the negative framework of privacy. The Lisbon Treaty has constituted a crucial further step in this process of constitutionalization, allowing the right to freedom of expression, private and family life, and the protection of personal data, as already enshrined in the Charter, to become binding vis-à-vis member states and European institutions, which can interfere with these rights only according to the Charter. Moreover, like the Convention, the Charter adds another important piece of the European constitutional puzzle by prohibiting the ‘destruction of any of the rights and freedoms recognized in this Charter or at their limitation to a greater extent than is provided for herein’”.

To be able to discuss and interpret the GDPR provisions and their applicability to TRUSTEE, as well as to the data processing activities envisaged to be conducted by the TRUSTEE service developers as well as TRUSTEE users, it is relevant to point out several definitions, included in the Glossary.

It is of high importance to understand the term of “personal data”, and the fact that not only direct, but also indirect identifiers may allow for identification of an individual, called in the GDPR nomenclature a data subject (for definition - see Glossary). Furthermore, it must be stressed that the GDPR distinguishes special categories of data (for definition - see the Glossary) and provides a stricter protection regime in this context. Article 9(1) GDPR states that processing of these types of data shall be, in principle, prohibited, unless any exception defined in Article 9(2) GDPR applies.

It should be stressed that even though, in the pilot use cases, there is the aim to use synthetic, fictitious or, if real, then only anonymised data (see D 1.5 – Data Management Plan), the TRUSTEE system has to be developed as the system ensuring appropriate and secure processing of personal data, especially the data considered as “special category”, such as above described data concerning health or genetic data, as well as personal data concerning minors, as these may be exchanged and explored by its future users, especially in the HEALTH and EDUCATION.

Considering that one of environments involved in TRUSTEE, for testing and validation of the solution is health domain, it is important to highlight and appropriately understand the term “data concerning health”. The definition included in the Glossary should be complemented by Recital 35⁸⁴, providing a wide interpretation of data concerning health, leading to a conclusion that all aspects of health should be considered to be part of the definition of personal data:

Personal data concerning health should include all data pertaining to the health status of a data subject which reveal information relating to the past, current or future physical or mental health status of the data subject. This includes information about the natural person collected in the course of the registration for, or the provision of, health care services as referred to in Directive 2011/24/EU of the European Parliament and of the Council (1) to that natural person; a number, symbol or particular assigned to a natural person to uniquely identify the natural person for health purposes; information derived from the testing or examination of a body part or bodily substance, including from genetic data and biological samples; and any information on, for example, a disease, disability, disease risk, medical history, clinical treatment or the physiological or biomedical state of the data subject independent of its source, for example from a physician or other health professional, a hospital, a medical device or an in vitro diagnostic test.

Another relevant environment in TRUSTEE which provides specific type of data, is EDUCATION. Although in principle protection of minors and children is considered as requiring additional, stricter approach, the data of a minor or of a child are not considered as special category of data according to the GDPR. Nevertheless, the processing based on Article 6(1)(a) GDPR (consent) should follow the requirement defined in Article 8 GDPR. In the case of the offer of information society services directly to a child, the processing of personal data of a child shall be lawful where the child is at least 16 years old. Where the child is below that age, such processing shall be lawful only if and to the extent that consent is given or authorized by the holder of parental responsibility over the child. Additionally, each Member State has the right to provide a lower age for those purposes provided that such lower age is not below 13 years.

⁸⁴ Although Recitals to EU laws are not in themselves legally binding like the operative provisions, nevertheless, where an EU law is vague, the recitals are an important means in interpreting the ambiguous provision. Recitals may be considered by the courts in the course of its purposive interpretation to EU law.

Anonymisation, Pseudonymisation, De-identification and Re-identification

In the context of ensuring data protection and privacy on TRUSTEE, it is important to consider and implement solutions which address limitation of the possibility of identification of the data subject. Such solutions decrease the risk of potential harm to the data subject in the situation where the data are compromised. The concepts relevant for TRUSTEE are anonymization and pseudonymisation of personal data, , as well as assessment whether the components included in TRUSTEE solution offer de-identification of the individuals, which would be made available by Data Providers and accessible by Consumers, and to which extent the re-identification of the individuals (data subjects) could be possible.

Information which does not relate to an identified or identifiable natural person or to personal data, rendered anonymous in a way that makes the data subject is not or no longer identifiable⁸⁵ is considered anonymous. Following this definition, anonymisation is process of removing personal identifiers, both direct and indirect, to achieve a stage in which a data subject cannot be re-identified by either the controller or another person, by using means reasonably likely to be used by them to re-identify an individual, whereby in the assessment of the “likelihood” the costs of identification, the time required for this and the available technology are all to be taken into account.⁸⁶ When data are fully anonymised and a person whose data are processed is no longer identifiable, the data fall outside of the scope of the GDPR.

Re-identification could be possible by: singling out, linkability of the data and inference.

An individual may be:

- directly identified from their name, address, postcode, telephone number, photograph or image, or some other unique personal characteristic.
- indirectly identifiable when the information is linked together with other sources of information, such as, for e.g., their place of work, salary, job title, postcode, location, the consumption of certain goods or the fact that they have a particular diagnosis or condition. This catalogue is not limited, and the aspect what is and what is not indirectly identifying personal information depends on a case-by-case assessment.

There are two different approaches which refer to the extent to which the re-identification should not be possible to state that the data is fully anonymised. The main approaches are relative and an absolute one. The latter stating that the anonymization should be in all instances irreversible, and even a data controller should not be able to re-identify a data subject after anonymisation process. Considering the possibility which AI models may give, fast developing technology and the increase of the possibility for inference and linkability, it is recommended to, by default, never consider raw datasets as anonymous. Further elaboration on pseudonymisation and anonymization can be found in the section concerning HE.

In contrary to anonymisation, pseudonymisation⁸⁷ techniques do not exempt controllers and processors from compliance with the GDPR. They do however help them meet their data protection obligations, related to the principles of ‘data minimisation’⁸⁸ and ‘storage limitation’⁸⁹, and application of ‘appropriate safeguards’, i.e. technical and organisational measures, relevant for maintenance of appropriate level of security of the data, as well as relevant for processing the data for scientific research purposes.⁹⁰ Unlikely anonymization in the absolute approach, pseudonymisation is considered a reversible process, as it is usually a process of replacing directly

⁸⁵ GDPR, Recital 26.

⁸⁶ GDPR, Recital 26.

⁸⁷ For the definition of „pseudonymisation“ see the Glossary.

⁸⁸ GDPR, Article 5(1)(c).

⁸⁹ GDPR, Article 5(1)(d).

⁹⁰ GDPR, Article 9(2)(j) and Article 89(1).

identifying information with indirectly identifying data. They may result from a cryptographic hash of the data of individuals, such as their IP address or user ID.

De-identification⁹¹ is a technique which may result in anonymisation or only in pseudonymisation of data – it depends on its scope. Anonymisation and de-identification are sometimes confused, i.e., de-identification in a form of simply removing personal identifiers (mostly, people's names) from data sets is understood also as anonymisation act. However, this type of act does not render the data anonymous (though, in some circumstance it may do so).

There are a number of techniques, which can be employed for the de-identification of data. These techniques are, e.g., Stripping-off names, addresses and other overt identifiers; Aggregating, or averaging, data across records; Broadening, coarsening or rounding data; Obscuring data clues but preserving event spacing by offsetting data by a random but uniform interval within each individual's records; Releasing only a fraction/part of the records from a set of data, which allows to achieve a result; Disruption.⁹² As indicated above, such techniques can either be considered as anonymisation or pseudonymisation, depending on the difficulty and likelihood of re-identification.

Re-identification is a means of linking anonymised data with other data sets such as the information available publicly in order to discover the identity of a data subject with whom the data relates.

PRINCIPLES OF PERSONAL DATA PROCESSING

Article 5 of the GDPR defines the principles which the entities processing personal data are abide by and must take into consideration already at the stage of designing a specific solution (such as TRUSTEE) to ensure high level of data protection, accountability and preservation of the rights of the data subject whose data will be processed via the solution. Additionally, they also need to always be kept in mind by the users of TRUSTEE, who may provide for further use (Data Providers), access and further process (Consumers) homomorphically encrypted personal data, or use the data in such a form for further training of the AI models (Model Providers/Consumers) made available on TRUSTEE. Each of the principles is shortly described below.

Lawfulness, Fairness and Transparency

The first principle requires that personal data shall be: *(a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');*⁹³

Lawfulness means that all processing of personal data must be conducted only (and at least for) one of the purposes set out by Article 6 GDPR. Special categories of personal data should be additionally processed based on one of the exceptions defined in Article 9(2) GDPR.

Fairness and Transparency: The principle of fairness of data processing is closely related to the transparency obligations. Transparency means “engendering trust in the processes which affect the citizen by enabling them to understand, and if necessary, challenge those processes”.⁹⁴ The controllers are obliged, where possible for them⁹⁵,

⁹¹ See also the Glossary.

⁹² ODI, Anonymisation and open data: An introduction to managing the risk of re-identification, (April 2019) accessed 4 August 2021. In bold – techniques which are relevant from TRUSTEE perspective.

⁹³ GDPR, Article 5(1)(a).

⁹⁴ Article 29 Data Protection Working Party, *Guidelines on Transparency under Regulation 2016/679* (WP 260 rev.01 2018), p. 4.

⁹⁵ Exceptions are defined in GDPR, Article 14(5).

to provide to the data subject information about data processing⁹⁶ in a concise, transparent, intelligible, easily accessible way and in a clear and plain language. In the context of application of the AI/ML models, this principle is particularly important, as the data subject must be made aware of the application of automated decision-making and shall have the right to refrain from being a subject to a decision based solely on automated data processing.

Purpose Limitation

The purpose limitation⁹⁷ ensures that any data processing is limited to specified purposes. The purpose shall be defined as precise as possible, and vague and general terms should be avoided in the process of defining it. It is possible to follow more than one purpose, but this should be revealed and communicated to the data subject, where possible. Nevertheless, there is a presumption included in the second half of this principle, which may play a relevant role for the Data Providers and Consumers of the data accessed via TRUSTEE. The presumption provides that further processing for scientific research purposes shall not be considered incompatible with the purpose for which personal data was originally collected. If such a case materialises, the controller (be it a Data provider or Consumer) who processes the data for such purposes, is not required to find a legal basis other than the one indicated for initial data processing, however, must ensure that technical and organisational measures, according to Article 89(1) GDPR (among them, especially pseudonymisation), are in place. With regards to the compatibility of purposes, another significant role in the secondary use of data plays a compatibility assessment, further defined in the section “Secondary use of personal data”.

Data Minimisation

When designing and developing TRUSTEE solution, it is particularly important that the principle of data minimisation⁹⁸ is considered at an early stage.

In line with this principle, controllers and processors (i.e., TRUSTEE users – Data providers and Consumers, as well as TRUSTEE system developers and providers) shall only process what is necessary, or in other words “relevant”, for the purpose of processing. In parallel, the data must be ‘adequate’ - though it must not exceed what is necessary for the original aim, it must still be enough so that it fulfils the purpose and is useful for the controller/processor.

“Controllers should be able to clearly explain and justify the need to collect and hold personal data, or consider using aggregated, anonymised or (when this provides sufficient protection) pseudonymised data for profiling.”⁹⁹

If the data can still be useful even if it is anonymised or masked, and the purposes for which they are processed fulfilled, then it should be considered that it is not necessary to have them in the form of personal data that allow for identification of data subjects. If the entity processing the data does not take it into account and processes the data in their original, unmodified shape, the data minimisation requirement cannot be acknowledged fulfilled.

There is a tension between the mentioned principle and the circumstance that, to ensure efficiency of the algorithms, they should be trained by and with the use of bigger amount of data, otherwise their performance may not be

⁹⁶ According to GDPR, Articles 12 to 14.

⁹⁷ GDPR, Article 5(1)(b) „personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes (‘purpose limitation’)“.

⁹⁸ GDPR, Article 5(1)(c) „personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (‘data minimisation’)“.

⁹⁹ Article 29 Data Protection Working Party, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* (WP 251 2017) at 11.

providing expected or appropriate and trustworthy results. Nevertheless, it needs to be emphasised that simple maximising the number of data used to train the models may have an impact on the spread of unlawful data collection activities. With relation to these doubts, one needs to understand that the data minimisation principle does not prohibit by default processing certain amount of data. It indicates that the data processing has to be always proven as necessary for its purposes. Therefore, the assessment providing if the use of specific sets of personal data to train AI models is “necessary” will always depend on the case, especially on the type of the AI solution and the goals for which it is used, as the factors determining data selection and inclusion. Moreover, controllers should set out limits that are sufficient to achieve the purpose of the processing, rather than using all available data.¹⁰⁰

Accuracy of processing

The principle of accuracy of processing defines that personal data shall be:

*(d) accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy')*¹⁰¹;

Data are inaccurate if they are misleading or incorrect.¹⁰² “Maintaining personal data to ensure accuracy should be built into regular process”.¹⁰³

Storage Limitation

The storage principle¹⁰⁴ requires that personal data should only be stored for a specified minimum amount of time.¹⁰⁵ Once the purpose of the processing is fulfilled, the data may no longer be stored except it is allowed by law. For scientific purposes, the GDPR acknowledges that personal data may be stored for longer periods, but controllers must apply technical and organisational measures to safeguard the rights and freedoms of the data subjects.

Confidentiality, Integrity and Data Security

According to integrity and confidentiality principle, personal data must be: *(f) processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality')*.¹⁰⁶

The maintenance of the security of personal data is a fundamental aspect of the right to data protection in accordance with Article 8 of the Charter. Both the controller (Data Provider and Consumer) and processor (TRUSTEE

¹⁰⁰ CIPL, Artificial Intelligence and Data protection: How the GDPR regulates AI (2020) (available here: <https://www.huntonprivacyblog.com/wp-content/uploads/sites/28/2020/03/cipl-hunton-andrews-kurth-legal-note-how-gdpr-regulates-ai-12-march-2020-1.pdf>), p. 13.

¹⁰¹ GDPR, Article 5(1)(d).

¹⁰² Peter Carey, Data Protection (fifth edn, Oxford 2018) 37.

¹⁰³ IT Governance Team, *EU General Data Protection Regulation (GDPR): An Implementation and Compliance Guide* (2nd edn, IT Governance Publishing 2018), p. 112.

¹⁰⁴ GDPR, Article 5(1)(e) „personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation')“.

¹⁰⁵ GDPR, Recital 39

¹⁰⁶ GDPR, Article 5(1)(f).

providers/system architecture developers and maintaining parties) need to consider the relevant risks associated with the processing of a particular type of data, and then to implement risk mitigation solutions in the form of appropriate technical and organisational measures to ensure security, confidentiality and integrity of the data.

When selecting measures protecting data, the sensitiveness of the type of the data needs to be considered. In the process of selection of required measures which the controller and processor must take to mitigate the risks, one can refer to Article 32 GDPR:

1. *Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate:*

- a) *the pseudonymisation and encryption of personal data.*
- b) *the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services.*
- c) *the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident.*
- d) *a process for regularly testing, assessing, and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.*

Selected measures which have been considered to be applied on TRUSTEE are described further in the section “Technical and organisational measures”, as well as in other sections of Part I “SOTA analysis” and Part V “Architecture” of this Deliverable. Further requirements related to the measures have been reflected in Part III of the Deliverable.

LEGAL BASIS FOR PROCESSING OF PERSONAL DATA

Processing of personal data should be lawful. This principle should be considered by the users: Data Providers and Consumers who would process personal data for their purposes, as well as by the TRUSTEE system developers and providers of the services which would be used by these entities, as these providers will also process the data of Data Providers and Consumers when offering the services (in this case it would be TRUSTEE system developers/providers).

Processing is lawful if (at least) one of the following is fulfilled:

- a. *a data subject has given consent to the processing of his/her personal data for one or more specific purposes¹⁰⁷;*
- b. *processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract¹⁰⁸;*
- c. *processing is necessary for compliance with a legal obligation to which the controller is subject¹⁰⁹;*
- d. *processing is necessary in order to protect the vital interests of the data subject or of another natural person¹¹⁰;*
- e. *processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller¹¹¹;*

¹⁰⁷ GDPR, Article 6(1)(a).

¹⁰⁸ GDPR, Article 6(1)(b).

¹⁰⁹ GDPR, Article 6(1)(c).

¹¹⁰ GDPR, Article 6(1)(d).

¹¹¹ GDPR, Article 6(1)(e).

- f. *processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.*¹¹²

The processing of special categories of personal data (such as, e.g., data concerning health) is in principle forbidden, unless one of the exemptions listed in Article 9(2) GDPR applies. It will be relevant for the Data Providers and Consumers of TRUSTEE, as well as TRUSTEE system developers/providers in a real setting, as such data may be theoretically processed via, or their processing facilitated by TRUSTEE, and mentioned actors will assume specific roles in the processing.

The selection of legal basis by Data Providers and Consumers appear to be completely dependent on them. Nevertheless, for the sake of trustworthiness, it would be recommended to require data providers and Consumers to indicate whether they do have a lawful ground for processing (encrypted) personal data. Nevertheless, it is not TRUSTEE responsibility, if these controllers cannot properly identify and justify it.

In relation to the data of the identification data/credentials of the users who would make use of TRUSTEE solution, for which TRUSTEE system developers/providers are controllers, it would be relevant that the providers indicate applicable legal basis for processing. At the moment, it seems that the basis defined in Article 6(1)(b), i.e. provision of a service, would be the most appropriate one.

DATA SUBJECT'S RIGHTS AND HOW TO ENABLE THEIR EXERCISE

In the context of providing and using complex technological solutions, which may be used for processing and sharing (accessing, transferring, disclosing) personal data, the data subjects' rights should be considered at all times.¹¹³ The controller is required¹¹⁴ to inform the data subjects about their rights and how to exercise them. This section on data subject's rights shall be considered as presentation of the rights which the data subjects (i.e. the individuals whose data would be provided by Data providers and further processed by Consumers, as well as TRUSTEE users' identification data) may have and which must be considered when designing the architecture of the TRUSTEE data space. A response procedure to the requests which may be raised by above-mentioned persons shall be developed and implemented by the TRUSTEE system developers.

Right to be informed - transparent information, communication and modalities for the exercise of the rights of the data subject¹¹⁵

What should be stressed in the first place is the fact that the requirements set by Article 12 to 14 are not a one-time requirement but should rather be considered throughout the whole processing life cycle.

Article 12 GDPR provides requirements on the manner in which the information about the data processing should be conveyed to the data subject. It is the controller's responsibility to ensure that the data subject is informed in a concise, transparent, intelligible and easily accessible way. By default the information should be "provided in writing, or by other means, including, where appropriate, by electronic means.

The data subject must be informed about the processing at the time when personal data is collected.¹¹⁶ If the controller plans to process the data for the purposes other than the one for which the data were collected, the data subject must be informed about that other purpose prior to the processing.¹¹⁷ It is relevant for Data Providers who may foresee to make their data accessible to the Consumers, and this activity would go beyond the initial purpose

¹¹² GDPR, Article 6(1)(f).

¹¹³ GDPR, Articles 15 to 23.

¹¹⁴ GDPR, Article 13 and 14.

¹¹⁵ GDPR, Article 12-14.

¹¹⁶ GDPR, Article 13(1) and (2).

¹¹⁷ GDPR, Article 13(3)

for which they collected the data from the individuals. It is also relevant for TRUSTEE system providers with regards to processing user credentials on TRUSTEE.

If the data is not collected directly from the data subject, but a controller obtained it from another controller, Article 14(3)(a) requires the new controller to provide the information within a “reasonable period”. This situation would take place in the case of Consumer on TRUSTEE, who would obtain access to the encrypted datasets being in the possession of the Data Provider.

However, in this context, it may be that the controller (in TRUSTEE setting this could be Consumer) is dismissed from the requirement to inform the data subject. Article 14(5) GDPR sets forth exceptions, where the controller is not required to provide the information to the data subject. This obligation can be lifted if a) the data subject already has the information, b) the provision of such information proves impossible or would involve a disproportionate effort – this applies especially in the situations where data are processed for the purposes of archiving for public interest, scientific or historical research or statistics. Moreover, also if fulfilment of this obligation is likely to become impossible or could seriously impair the achievement of the processing objectives.

In mentioned cases, the controller shall take appropriate measures to ensure the protection of the data subject’s rights and interests, by making the relevant information publicly available.

Right of Access¹¹⁸

Individuals have the right to be informed whether or not personal data concerning them are processed. The controller shall provide all information required¹¹⁹ (e.g., purpose, categories concerned, recipients) including a copy of personal data. If data are transferred to a third country the data subject must be informed about the safeguards identified in Article 46 GDPR.

If the data subject makes a request by electronic means, the answer, including requested information, shall also be provided electronically, however, the data subject may request to receive the information in another form.

Right to Rectification¹²⁰

In line with the principle of accuracy, the controller must rectify all inaccurate data from the data subject, including completing incomplete data. In this context, Data Providers should have the possibility to reply to this type of request and rectify the data of the individuals whose data are made available by Data Providers. Additionally, the TRUSTEE system developers/providers should have the possibility to rectify all inaccurate personal data of the TRUSTEE users (Data Providers, Consumers, Model Providers, and Developers).

Right to Erasure¹²¹

The data subjects (either individuals whose data are made available by a Data Provider or the users of TRUSTEE – Data Providers, Model Providers, Consumers, Developers) may request deletion of their data without undue delay. Therefore, if:

- Their personal data are no longer necessary to be processed with regards to the purposes for which they were collected.
- The data subject withdraws consent for processing the data, and there is no other legal basis that would apply.
- The data subject objects to the processing and no overriding legitimate grounds for the processing apply.
- The personal data have been processed in an unlawful manner.
- Erasure of the data is required by Union or Member State law.

¹¹⁸ GDPR, Article 15.

¹¹⁹ GDPR Article 15(1).

¹²⁰ GDPR, Article 16.

¹²¹ GDPR, Article 17.

- The data have been collected in relation to the offer of information society services (e.g., marketing purposes).
- the controller has to delete the data.¹²²

There are several exceptions where the right to erasure does not apply. It is very much dependent on the case, whether the data of an individual would need to be deleted by the Data Provider, or data of the user would need to be deleted by TRUSTEE system developers and providers. The data can be kept if the processing is necessary:

- for exercising the right of freedom of expression and information.
- for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.
- with regard to processing of special categories of personal data - for reasons of public interest in the area of public health (further defined in Article 9(2)(h) and (i) and 9(3) GDPR);
- for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) GDPR in so far as the right to erasure is likely to render impossible or seriously impair the achievement of the objectives of that processing; or
- for the establishment, exercise or defence of legal claims.¹²³

If none of the exceptions defined in Article 17 GDPR applies, the Data Providers, where applicable – Consumers, as well as TRUSTEE system developers should have the possibility to erase data as requested a) by the data subjects whose data are made available via TRUSTEE by their users or b) TRUSTEE users themselves with regards to their identification data.

Right to Restriction of Processing¹²⁴

The data subject should have the right to restrict processing of his or her data by the controller where one of the following cases applies:

1. *the accuracy of the personal data is contested by the data subject, for a period enabling the controller to verify the accuracy of the personal data.*
2. *the processing is unlawful, and the data subject opposes the erasure of the personal data and requests the restriction of their use instead;*
3. *the controller no longer needs the personal data for the purposes of the processing, but they are required by the data subject for the establishment, exercise or defence of legal claims.*
4. *the data subject has objected to processing pursuant to Article 21(1) pending the verification whether the legitimate grounds of the controller override those of the data subject.*¹²⁵

Where the processing has been restricted based on one of the provisions indicated above, any processing, is allowed only with the data subject's consent or it is necessary for the establishment and exercise of legal claims.¹²⁶ It does not apply to the storage of data.

In the context of TRUSTEE solution, it is therefore important for TRUSTEE system developers and providers, as well as users of TRUSTEE who are making the data available or will process the data in their encrypted format, to

¹²² GDPR, Article 17(1).

¹²³ GDPR, Article 17(5).

¹²⁴ GDPR, Article 18.

¹²⁵ GDPR, Article 18(1).

¹²⁶ GDPR, Article 18(2).

assess the situation based on a specific case, whenever such a request is made by a data subject. A protocol should be implemented to address whether the processing of data subject's data fulfils one of the requirements set forth above and if so, the controller has a twofold solution: should either ask for consent or limit the processing just to storage and not undertake any further processing.

Right to Data Portability¹²⁷

The right to data portability is one of the pillars of the free flow of personal data and reflection of data subjects autonomy and right to dispose over their data. Individuals have the right to receive from a controller the personal data they have provided, , in a structured, commonly used and machine-readable format¹²⁸ and request their transmission to another controller, where technically feasible – directly¹²⁹ - without hindrance where:

- a) the processing is carried by automated means,
- b) where the processing is based on consent and
- c) where the processing is based on a contract pursuant to Article 6(1)(b).¹³⁰

The right to data portability is subject to some limitations - for example, it does not apply when the performance of a task carried out in the public interest requires that the data processing takes place.¹³¹ The exercise of the right should not be considered as taking priority over the right to erasure and shall not affect the rights and freedoms of others in an adverse manner.¹³²

Right to Object¹³³

According to Article 21(1) GDPR, data subjects have the right to object to the processing of their personal data when the processing is:

- a) Article 6(1)(e): necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.
- b) Article 6(1)(f): necessary for the purpose of the legitimate interest pursued by the controller or a third party.

If the controller is able to demonstrate compelling legitimate reasons for the processing that override the rights of the mentioned data subject or for the defense of a legal claim, the right to object might be limited.¹³⁴

When data subjects exercise their right and raise an objection, the processing of their personal data for a specific purpose that they objected to becomes unlawful. Where personal data are processed for scientific research purposes, data subjects can object to such processing unless it is necessary for the performance of a task carried out for reasons of public interest.¹³⁵

National Derogations from the Rights of a Data Subject

The GDPR in its Article 85 and Article 89 foresees the possibility for Member States to implement derogations to certain rights defined above, when personal data are processed for:

¹²⁷ GDPR, Article 20.

¹²⁸ GDPR, Article 20(1).

¹²⁹ GDPR, Article 20(2).

¹³⁰ GDPR, Article 20(1).

¹³¹ GDPR, Article 20(3).

¹³² GDPR, Article 20(4).

¹³³ GDPR, Article 21.

¹³⁴ GDPR, Article 21(1).

¹³⁵ GDPR, Article 21(6).

- a) journalistic purposes and the purposes of academic, artistic or literary expression. In this context, national legislation can provide for derogations or exemptions from all rights of the data subject.¹³⁶
- b) archiving, statistical, scientific and historic purposes. In these cases, the rights to: access, rectification, restriction of processing and object can be limited by national law.

TRUSTEE solution is designed to operate in various countries across Europe. This requires further analysis of all national laws in which the TRUSTEE solution is planned to be made available. At the moment, the focus has been put on the legislation at the European level, nevertheless, in the case it becomes relevant, appropriate assessment of national jurisdictions will be conducted, and relevant outcomes presented in the deliverables relevant for specific pilots (Deliverables of WP5).

ROLES WITH RESPECT TO THE DATA PROTECTION

In the course of the design and development of the TRUSTEE solution it is relevant to consider and identify the roles which different actors: TRUSTEE system developers and providers, and users such as Data Providers, (Data) Consumers, would assume in line with the data protection regime. This plays a significant role with regards to establishing relevant formal and technical solutions for the future use of TRUSTEE: in a real world and with datasets which may represent personal data.

A controller¹³⁷ is responsible for processing of personal data according to the GDPR. The controller is required to take into account all legal and regulatory aspects related to the data processing activities, first and foremost consider the risks that are related to processing and its possible impact on the rights and freedoms of the individual whose data are processed. Additionally, the controller is obliged to adopt safeguards to protect processed personal data and ensure and respond to the data subject rights.

In the relationship between identified TRUSTEE users (i.e., Data Providers and (Data) Consumers), it appears that both figures would become separate controllers of the data accessed by the Consumer at Data provider's premises, via TRUSTEE, and further processed by the Consumer. In this constellation, Data Provider would be responsible for appropriate provision of the data and ascertaining that making them available to the Consumer, in an encrypted form, stays in the frame of the purposes for processing, which a data subject whose data are accessed is aware of and for which the data are collected for. In case it is not fulfilled, Data providers would need to conduct a compatibility assessment (described further in this Section), to see if the processing is compatible with the initial purpose, and if not – as a data subject for consent or make sure that there is another legal basis that allows for the processing. In any case, however, Data provider would be required to inform a data subject about this new purpose.

Consumers, who would access homomorphically encrypted data and process them in that form would set out their own, separate processing purposes, , not necessarily the ones set out by the initial data controllers who collected the data, and hence, would become separate controllers, obliged to fulfil all relevant obligations stemming from the GDPR and applicable to controllers on their own.

Another relevant role from data protection laws to be considered in TRUSTEE is processor¹³⁸. The processor is allowed to process personal data only according to explicit instructions from the controller, unless required by Union or Member States law¹³⁹, and only if the processor is able to ensure sufficient guarantees that technical and organisational measures appropriate for ensuring data protection will be applied.¹⁴⁰ Processors are prohibited to

¹³⁶ GDPR, Article 85(2).

¹³⁷ GDPR, Article 4(7), see also the definition in Glossary.

¹³⁸ Definition can be found in the GDPR, Article 4(8) and in the Glossary.

¹³⁹ GDPR, Article 29.

¹⁴⁰ GDPR, Article 28(1).

engage another processor (sub-processor) unless based on an authorisation expressed (in written or an electronic form) by the controller.¹⁴¹

The controller's and the processor's relationship should be governed either by a contract (concluded in writing, incl. an electronic form) or other legal act. The contract must address at least the following points: the purpose and subject-matter, duration of required processing, nature, type of personal data and categories of the data subject, and this all additionally to the rights of the controller.

Considering the TRUSTEE solution and its components, TRUSTEE system developers and providers, allowing the users for making their data available in a secure manner by applying TRUSTEE means, might assume the role of the data processor if the system allows for processing of real datasets (e.g. making accessible between Data providers and Consumers and allowing for processing the data on Data provider's premises thanks to the federated learning approach) where the datasets may be considered as pseudonymised personal data (due to the application of homomorphic encryption or other pseudonymisation techniques).

A third party is any natural or legal person, public authority, agency, or any other body other than the data subject, the controller, the processor and the person who, under the direct authority of the controller or the processor, are authorised to process the data.¹⁴²

To sum up, having considered TRUSTEE system architecture and the relationships between different actors involved in TRUSTEE processes, there are several relationships, which should be considered and, based on the initial assessment, reflected in appropriate agreements:

- Relationship between TRUSTEE users: Data Providers and Consumers.
This relationship could be regulated agreements, which could be signed via TRUSTEE (thanks to the tool developed in T4.6 – Security and Trust Manager (STM)). . However, the agreement would be considered as signed between two separate controllers, and would de facto be a “data access” agreement, and it would regulate the conditions according to which Consumer could access and process encrypted data from Data Provider via TRUSTEE. The agreements would not involve processors.
- Relationship between the TRUSTEE system developers/providers, and respective users (Data Providers and Consumers) of TRUSTEE – and this needs to be understood twofold:
 - *Where the subject to the agreement are the data made available by Data Providers with the use of TRUSTEE components in an encrypted form to the Consumers for the processing.*

Having analysed the relationship between the users of TRUSTEE and the TRUSTEE system developers/providers, it seems that the former group would represent a type of “controllers” of the (personal, encrypted) data they share, i.e. make accessible / disclose (Data Providers), or they access (Consumers), who processes the data for their own purposes using the means being part of TRUSTEE. The TRUSTEE system developers/providers would be in this context processors towards the users, as they would, when providing the structure for the data sharing process encrypted data on behalf of the users, using the means provided by TRUSTEE, which should be also indicated and described in the Terms of Use of the system, compliant with the Article 32 of the GDPR.

- *Where the subject to the agreement are the user's data, such as log in/identification data/credentials, provided by the users onto TRUSTEE to be able to use the solution for data exchange and manipulation.*

In this case parties providing TRUSTEE infrastructure would be controllers of the users' data, and they would be processing them for the purposes of providing TRUSTEE service to the users.

¹⁴¹ GDPR, Article 28(1).

¹⁴² GDPR, Article 4(10).

DATA PROTECTION BY DESIGN AND BY DEFAULT

In the regime of data protection, it is of crucial to implement data protection by design¹⁴³ approach, which incorporates privacy and data protection at the stage of the design and development of a system or a process— in the case of case the architecture of the technological solution that TRUSTEE represents. To comply with this requirement, TRUSTEE system developers should consider the safeguards defined in the GDPR in the design, development and operational phase of the TRUSTEE platform.¹⁴⁴ The architecture of TRUSTEE system must involve appropriate technical and organisational measures designed to address the GDPR principles.¹⁴⁵ In the process of selecting the measures, risk management approach should be applied.

Data protection by default is another obligation imposed on the controllers and processors¹⁴⁶. It requires implementation of default settings which ensure that processed data will be just personal data necessary for the fulfilment of specific processing purposes. TRUSTEE system should favour by default a protective setting ensuring the highest level of protection when no action is taken by the data subject.¹⁴⁷ Additionally, it should be also considered that the users or the data subjects whose data are provided by the users might wish to exercise their rights. Therefore, where applicable, the TRUSTEE solution should enable smooth and timely reaction to their requests.

AUTOMATED INDIVIDUAL DECISION-MAKING, INCLUDING PROFILING

In accordance with Article 22(1) GDPR, data subjects have the “right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects or significantly affects him or her”.¹⁴⁸ This should be considered as a general prohibition for automated individual decisions that are made without human involvement or intervention in the decision-making process.¹⁴⁹

However, the GDPR introduces three possible exceptions from this general prohibition:

- a) is necessary for entering into, or performance of, a contract between the data subject and a data controller.
- b) is authorised by Union or Member State law to which the controller is subject, and which also lays down suitable measures to safeguard the data subject’s rights and freedoms and legitimate interests; or
- c) is based on the data subject’s explicit consent.¹⁵⁰

In order to fully fall under restriction of Article 22(1), the decision must ‘solely’ be based on automated processing. The term ‘based solely’ translates into a situation in which no human is involved in the decision process. It is therefore important that whenever a decision concerning a data subject is based on the result of processing performed with the use of AI models available on TRUSTEE, the users (Consumers or Data Providers) who employ

¹⁴³ GDPR, Article 25(1).

¹⁴⁴ European Union Agency for Network and Information Security (ENISA), Privacy and Security in Personal Data Clouds (Final Report, Public, November 2016), p. 16.

¹⁴⁵ European Union Agency for Network and Information Security (ENISA), Privacy and Security in Personal Data Clouds (Final Report, Public, November 2016), p. 6.

¹⁴⁶ GDPR, Article 25(2).

¹⁴⁷ European Data Protection Supervisor, Preliminary Opinion on Privacy by Design (Opinion 5/2018, May 2018), p. 7.

¹⁴⁸ GDPR, Article 22(1).

¹⁴⁹ Article 29 Data Protection Working Party, 'Guidelines on Automatic individual decision-making and Profiling for the purposes of Regulation 2016/679' (2018) at 20.

¹⁵⁰ GDPR, Article 22(2).

the AI models will participate in the decision-making process and will have the possibility to intervene and should be capable of discovering and recovering unfair outcomes or discriminations.¹⁵¹

TECHNICAL AND ORGANISATIONAL MEASURES

One of the core elements and aims of the TRUSTEE solution is the selection (or design) and implementation of the measures ensuring data confidentiality, integrity and security, according to the principle defined in Article 5(1)(f).

According to Article 32 GDPR, controller and processor have an obligation to implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk to the rights and freedoms of natural person in the processing of personal data. In the context of selecting proper security measures, the factors such as the state of the art, the costs of implementation and the nature, scope context and purposes of processing have to be taken into account in the assessment of the appropriateness of the security level.¹⁵²

The Article 32(1) GDPR provides a number of specific security measures which may be implemented, namely:

- a. *Pseudonymisation and encryption of personal data.*
- b. *Measures to ensure the security of IT employed in the data processing, i.e., ensuring its ongoing confidentiality, integrity, availability and resilience;*
- c. *Incident response capabilities to restore the availability and access to personal data in a timely manner.*
- d. *A regular assessment and evaluation of the effectiveness of the implemented technical and organisational security measures.*

This catalogue is not closed; hence controllers and processors can go beyond it and seek other, perhaps evedvanced, dvanced methods of protecting the data and securing data processing activities.

Though during most of the pilot use cases only fictitious data will be used, for the future exploitation of the TRUSTEE solution and with the consideration that real data, especially personal data, can be processed in TRUSTEE, TRUSTEE should implement and test necessary technical and organisational measures that can ensure an appropriate level of security of personal data.

Additionally, it should be considered whether any third party such as an external cloud provider would be involved in the activities or maintenance of TRUSTEE. Considering the current stage of the project, it is not defined yet, who would host the service. This should be subject to further discussions between TRUSTEE system developers, whether the service will be developed and maintained on a local cloud of one or several of the developers, or an external cloud infrastructure provided by a third party. In the latter case, appropriate arrangements also with the cloud service providers should be made.

With regards to the integrity¹⁵³ requirement, when defining the TRUSTEE architecture, the processing of personal data, i.e. making data available /disclosing the data by Data Provider to the Consumer, and using the data, which were encrypted with TRUSTEE support, by Consumer, has to remain unaltered. To help ascertain that, cryptographic SSI authentication mechanisms could be considered to be employed to detect alterations to personal data.

¹⁵¹ EDPB, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default (13 November 2019) https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf , accessed 28.12.2022.

¹⁵² GDPR, Article 32(1).

¹⁵³ See the Glossary for the definition.

To address confidentiality requirement, the encryption of data “in transit” as well as “at rest” is recommended to be considered, where feasible and operationally practicable. Although encryption usually does not suffice to render personal data anonymous, it definitely improves the level of security and confidentiality by limiting the possibility of accessing the data by unauthorized parties. Additionally, technical measures ensuring confidentiality such as appropriate authorisation and strong authentication mechanisms shall be put in place¹⁵⁴, to allow for preventing, unauthorized access to data, identity theft, unlawful processing and data leakage.

TRUSTEE offers a comprehensive set of components which ensure privacy and security by design and appropriate level of data protection, further described in Part V Architecture. The below sections provide considerations on two core technologies which TRUSTEE builds upon and which shall allow for cutting-edge approach to data security and privacy: HE and federated learning, and their interplay with the data protection regulations and their main concepts. The sections are complementary to Sections “Homomorphic-capable self-sovereign framework” and “Privacy and Security in federated Learning” and present initial considerations on the topic. Further developments and advancements in assessment of the technologies in TRUSTEE environment will be reflected in the updates to the subject Deliverable. The requirements stemming from Article 32, included in Part III of this Deliverable, have been also complemented by relevant guidelines on personal data security in the development phase, issued by the French Data Protection Authority (CNIL)¹⁵⁵.

HOMOMORPHIC ENCRYPTION IN TRUSTEE

HE is considered as one of the encryption types, which addresses Article 32(1)(a) GDPR.

It is a privacy enhancing technique that allows for operations to be carried out on encrypted data. The result of the operation is equivalent to the results which would be obtained working directly with original, unencrypted information.¹⁵⁶ It is suitable for cases where a controller collaborates with an entity, who is required or allowed to carry out a processing activity, and the controller wishes to ensure that the entity cannot directly access the data, i.e. see their exact content.¹⁵⁷ It focuses on hiding the data by protecting the privacy of data subjects without affecting accuracy and utility of the data.¹⁵⁸

It is recognized by various Data Protection Authorities and cybersecurity, such as:

- Spanish Data Protection Authority (AEDP) considers HE as one of advanced TOMs (technical and organisational measures) corresponding to Privacy by Default requirement¹⁵⁹
- Hellenic Data Protection Authority recognises HE as one of the tools used for pseudonymisation purposes¹⁶⁰
- ICO (the UK Data Protection Authority) perceives HE as one of the Privacy Enhancing Techniques focusing on hiding /shielding the data, providing strong security and confidentiality¹⁶¹
- ENISA perceives HE not as the tool which could lead to anonymisation *per se*, but as a “data masking” technique which could lead to “pseudonymisation”¹⁶².

¹⁵⁴ Article 29 Data Protection Working Party, Opinion 5/2012 on Cloud Computing (WP 196 2012), p. 14.

¹⁵⁵ [GDPR developer's guide | CNIL](#).

¹⁵⁶ [Encryption and Privacy III: Homomorphic encryption | AEPD](#).

¹⁵⁷ Ibid.

¹⁵⁸ ICO, [Chapter 5: Privacy-enhancing technologies \(PETs\), Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance](#), p. 6.

¹⁵⁹ See [Encryption and Privacy III: Homomorphic encryption | AEPD](#).

¹⁶⁰ See [Cryptography at the service of pseudonymisation \(europa.eu\)](#)

¹⁶¹ See - [Chapter 5: Privacy-enhancing technologies \(PETs\), Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance](#); [Chapter 2: How do we ensure anonymisation is effective? \(ico.org.uk\)](#)

¹⁶² Further on it: ENISA, Data Protection Engineering. From Theory to Practice, January 2022.

According to the above mentioned ICO, in the context of processing of personal data¹⁶³, this activity of processing homomorphically encrypted data likely counts as “consultation” and “use” of personal data, with the encryption representing “adaptation or alteration” of the data. Hence, the data is still personal data, just treated in a particular way that enhances the privacy of those to whom it relates.¹⁶⁴

It ensures security and confidentiality by minimizing the risk associated with the data breaches, when they occur, as personal data remains encrypted all the time – at rest, in transit and during computation activities. The data becomes then unintelligible for an intruder and thanks to that the risks to the data subjects decrease.¹⁶⁵ It increases the level of accuracy, by assuring (at least to some extent) that the outcome of a computation activity is the same as the activity performed on unencrypted data. However, to ensure that, it must be confirmed that the inputs are correct before encrypting them. Otherwise, the result may be different from the processing performed on unencrypted dataset.

It is of utmost importance, from the data protection and privacy perspective, to ensure appropriate choice of the right algorithm, right key size, right software and of keeping the key secure.¹⁶⁶ It is also crucial to apply TOMs which ensure the security of the secret key, as well as necessary to put in place a process which will generate a new key right after the original one has been compromised.¹⁶⁷

Moreover, the effectiveness of the solution should be constantly monitored and improved, taking into account the development of decryption technologies.

The latest ruling of the General Court of the EU¹⁶⁸ included the reasoning that the assessment whether the data is anonymous or only pseudonymous depends on the perspective and in whose hands the data is and what powers the entity accessing and processing the data has that would enable to re-identify a data subject. The court ruled that it should be always assessed on a case-by-case basis whether controllers (and in the case of TRUSTEE Data provider and Consumer) would have legal means available to them, which in practice would enable to access the additional information necessary to re-identify a data subject. In other words, the assessment of feasibility and effort needs to be carried out from the recipient’s perspective, and it should be taken into account that sometimes recipients of the data are not provided with the key that would allow to re-identify the person behind the processed data, and hence – the data would be non-personal for this recipient.

This may have an important impact on the further assessment in the TRUSTEE context whether a Consumer and TRUSTEE system developers could be considered controller and/or processors, or if they could argue that due to the fact that they do not have access to the key, they process non-personal data and, in this structure, only the Data provider can be considered as an entity processing personal data.

Since the ruling is not definite, and can be appealed, further considerations on it will be presented in D2.2 and D2.3.

FEDERATED LEARNING IN TRUSTEE

¹⁶³ Processing, as defined in Article 4(2) GDPR, ‘means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction’;

¹⁶⁴ ICO, [Chapter 5: Privacy-enhancing technologies \(PETs\), Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance](#), p. 13.

¹⁶⁵ Ibid.

¹⁶⁶ Ibid., p. 14.

¹⁶⁷ Ibid.

¹⁶⁸ Case T-557/20, SRB v EDPS. This ruling can be appealed with the European Court of Justice.

Considerations regarding federated learning and processing personal data from the GDPR perspective:

Federated learning technology allows to ensure that processing is done in a GDPR compliant manner and addresses the “privacy by design” principle under Article 25 GDPR, which necessitates the implementation of privacy enhancing technologies to preserve the privacy and safety of personal data and also to abide by the data protection principles envisaged under the GDPR. In the context of federated learning, the emphasis should be given to the following principles:

- clear identification of purpose, objectives and limitations (purpose limitation);
- controlling the accuracy of data (accuracy);
- adjusting the “amount of data to purpose and training phases”¹⁶⁹ (data minimization).

Concerning the data minimisation principle, though it is argued that federated learning ensures data minimisation and separation of the processing of personal data to an important extent as the amount of raw personal data shared with other parties is minimised¹⁷⁰ and the data locally centred are not revealed to other parties or to a central source in its raw format, the principle that personal data used for the training should be “adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed”¹⁷¹ should still apply. This is due to the assertion that the aim of these models is to be able to infer insights from as much data as possible without sharing the data with other parties.¹⁷² Thus, during the training of the federated data model, only data that is fitting and adequate for the purpose of developing the model should be used. This could also minimise, in the phase where local/client models are merged with the central model at server level, the risk of inference of information used in training the models locally.

The biggest advantage of federated learning models in terms of protection of privacy and data security is that the raw data used in the training does not have to be transferred to the central or other servers, as the model is trained separately at each device/local server with the data held within. Thus, less communication in this sense is described to mean less risk of data leakage.¹⁷³

However, the sole implementation of a federated data model still does not make the TRUSTEE users exempt from the application of GDPR. The literature does not find the singular implementation of federated learning models adequate to preserve privacy and security and completely avoid data leakage. Though the threat of direct leakage is proclaimed to be mostly avoided, it is discussed in literature that “personal information can be stealthily extracted from local training parameters” despite implementing a federated learning model, which puts the models within the scope of the GDPR.¹⁷⁴ Therefore, the development of the federated learning models should still consider the effects of the insertion of personal data into the TRUSTEE ecosystem.

As a first note, if the raw training data, on which the operations of the federated data models are performed, can be categorized as personal data, the provisions of the GDPR shall become applicable. However, the GDPR, within its

¹⁶⁹ Sophie Stalla-Bourdillon, Alfred Rossi, Gabriela Zafir-Fortuna, “Data Protection by Process How to Operationalize Data Protection by Design for Machine Learning”, December 2019, access from <https://fpf.org/wp-content/uploads/2019/12/WhitePaper_DataProtectionByProcess.pdf>.

¹⁷⁰ Data protection by design in AI? The case of federated learning by Stephanie Rossello, Roberto Díaz Morales and Luis Muñoz-González, 8.

¹⁷¹ GDPR, Article 5(1)(c).

¹⁷² Privacy preservation in federated learning: An insightful survey from the GDPR perspective, Nguyen Truong, Kai Suna, Siyao Wang, Florian Guittona, YiKe Guo, Computers and Security 110 (2021), 5.

¹⁷³ Sheng Shen, Tianqing Zhu, Di Wu, Wei Wang, Wanlei Zhou, “From distributed machine learning to federated learning: In the view of data privacy and security”, Concurrency Computat Pract Exper. 2022 (34), p.6.

¹⁷⁴ Truong, Suna, Wang, Guittona, and Guo (n 4) 2.

material scope, could also encompass “the operations performed on the model updates derived from the ‘raw’ training data”.¹⁷⁵ If the models can reveal information regarding the underlying personal data used for training the models, by “all the means reasonably likely to be used to identify the natural person directly or indirectly”¹⁷⁶, then it is argued that the GDPR would become applicable.¹⁷⁷

Researchers in the field assert that data leakage could be probable at two different stages of federated learning: either at the local device level or the central server level:

Local device level threats: Any data leakage from the local devices could be the result of inadequate implementation of privacy preserving and enhancing mechanisms locally to protect the training data used locally as well as working with non-trusted or comprised partners in training the central model in the federated learning process.

Central server level threats: The potential data leakage problem at the central server level refers to the possibility that the parameters with which each local model is updated and then shared with the central server, could also lead to leakage during transmission¹⁷⁸ to “restore client data” via “the ability of the training model to remember the characteristics of the data”.¹⁷⁹ Another problem elaborated on in the literature is the inference problem¹⁸⁰, where through the sharing of models rather than raw data, adversaries can still infer certain characteristics of the training data used at local sites.¹⁸¹ There could be an inference threat by the merging of parameters coming from multiple local devices.

Overall, as the European Data Protection Supervisor (“EDPS”) also indicates, with the possibility of some types of attacks (i.e. model poisoning) due to “the distributed nature of federated learning”, the inadequate protection of “classic defence mechanisms” in these settings require for the testing and implementation of “ad hoc defence methods.”¹⁸² It is not only the EDPS who advises this, but the necessity to execute other privacy enhancing

¹⁷⁵Rossello, Díaz Morales, and Muñoz-González (n 2) 9

¹⁷⁶ GDPR, Recital 26.

¹⁷⁷ Rossello, Díaz Morales, and Muñoz-González (n 2) 9. Please also see Jessica Chia Liu, Jack Goetz, Srijan Sen, Ambuj Tewari, “Learning From Others Without Sacrificing Privacy: Simulation Comparing Centralized and Federated Machine Learning on Mobile Health Data”, JMIR Mhealth Uhealth 2021, 9(3), p.5: “Even if no user device data are transmitted to the server, it may be possible to infer information about the data used to train the model, given sufficient access to the model itself, including the reconstruction of specific data points used to train the model...complementary forms of privacy may be required.”

¹⁷⁸ Rossello, Díaz Morales, and Muñoz-González (n 2) 4.

¹⁷⁹ Chuanyin Wang, Cunqing Ma, Min Li, Neng Gao, Yifei Zhang, Zhuoxiang Shen, “Protecting Data Privacy in Federated Learning Combining Differential Privacy and Weak Encryption” in Science of Cyber Security Third International Conference, SciSec 2021 Virtual Event, August 13–15, 2021 Revised Selected Papers, Springer (2021), p. 95, 96. See also Matthew G. Crowson, Dana Moukheiber, Aldo Robles Arevalo, Barbara D. Lam, Sreekar Mantena, Aakanksha Rana, Deborah Goss, David W. Bates, Leo Anthony Celi, “Federated learning applications for biomedical data”, PLOS Digital Health (2022) p.11: “While FL allows ML model building without raw data collection from multiple institutions, the possibility for an adversary to learn about or memorize sensitive user data by simply tweaking the input datasets and probing the output of the algorithm exists... ,note that there is a crucial tradeoff between the convergence of the ML models during training and privacy protection levels as better convergence comes with lower privacy.”

¹⁸⁰ This is even described in literature as one possible direction of research: “Therefore, one future direction of research is to investigate how to prevent data leaks due to inference during the training process.” Please see: Sheng Shen, Tianqing Zhu, Di Wu, Wei Wang, Wanlei Zhou, “From distributed machine learning to federated learning: In the view of data privacy and security”, Concurrency Computat Pract Exper. 2022 (34), p.16.

¹⁸¹ Sheng Shen, Tianqing Zhu, Di Wu, Wei Wang, Wanlei Zhou, “From distributed machine learning to federated learning: In the view of data privacy and security”, Concurrency Computat Pract Exper. 2022 (34), p. 2.

¹⁸² [Federated Learning | European Data Protection Supervisor \(europa.eu\)](#) It should be reminded that the entity running the central server for the aggregation of different local models and updates would need to ensure that the

technologies, such as the application of HE, is emphasised by many others. TRUSTEE system developers need to explore well how the federated data model functions in a homomorphically encrypted framework and should assess the effectiveness of both technologies in parallel. They should also be prepared and testing against possible attacks to the federated model to be developed both at local and central level to develop system security.

Considerations regarding non-personal data:

Not only personal data but also non-personal data, specifically requiring a certain degree of confidentiality, is foreseen to be exchanged via the TRUSTEE framework.

Therefore, for the training and development of the federated data model in TRUSTEE via the use of non-personal data, the users involved in these activities should implement strong security mechanisms concerning primarily but not limited to:

- the persons having access to the training data used and the model trained from their institutions: only those people who need to be involved and have access should be granted access;
- the methods/schemas for granting access: identification and authentication mechanisms should be put in place.

Each party making use of TRUSTEE should make sure that they are not under any contractual obligation limiting their use of the datasets, which they would like to make accessible via TRUSTEE, due to i.e. confidentiality reasons. The users will be individually liable for any breach of contract in that respect.

To sum up above mentioned considerations - the federated learning model in TRUSTEE should be designed with having in mind that in the future, personal data or any other confidential or sensitive data may be brought to the TRUSTEE solution to be shared between users (Data Providers and Consumers). Thus, the model needs to be developed as having the capability to act as a privacy enhancing technology when faced with personal data or confidential data during its exploitation. In order to ensure the confidentiality and integrity of such data exchanged via the use of the federated model developed and also to ensure that the TRUSTEE system design is robust and privacy preserving, the aim should be the implementation of privacy preserving techniques not just to the global/central model but also at each local model training the algorithm with raw data held at local servers. Though homomorphic encryption, as a cryptographic method, is mentioned as a privacy enhancing technique in the literature¹⁸³, due to the multiplication of sites within such federated models, each site needs to be diligent with the security measures put in place for the protection of personal data or any other sensitive data used. The users should have a common approach in requiring strong mechanisms to be implemented at each site, such as the implementation of sturdy access control mechanisms and verification/identification systems to assure the participation of trusted institutions.¹⁸⁴

updated models are processed in a secure manner per Article 32 GDPR, which envisages the security of processing and appropriate safeguards to be applied by data controllers and processors.

¹⁸³ Please see Sheng Shen, Tianqing Zhu, Di Wu, Wei Wang, Wanlei Zhou, “From distributed machine learning to federated learning: In the view of data privacy and security”, *Concurrency Computat Pract Exper.* 2022 (34), p.9. Some studies suggest “multi-key homomorphic encryption protocol” as an effective measure “to design novel privacy-preserving federated learning scheme”, in which “model updates are encrypted via an aggregated public key before sharing with a server for aggregation.” Each model participating in aggregated model does not share the same encryption and decryption keys in order to avoid information to leak among different participants. Please see Jing Ma, Si-Ahmed Naas, Stephan Sigg, Xixiang Lyu, “Privacy-preserving federated learning based on multi-key homomorphic encryption”, *Int J Intell Syst.* 2022 (37), 5880–5901, p. 5880 and 5884. See also Truong, Suna, Wang, Guittona, and Guo (n 4) 16.

¹⁸⁴ Ons Aouedi, Alessio Sacco, Kandaraj Piamrat, Guido Marchetto, “Handling Privacy-Sensitive Medical Data with Federated Learning: Challenges and Future Directions”, *IEEE JOURNAL OF BIOMEDICAL AND HEALTH*

Furthermore, with respect to personal data, there exists a potential risk of identifying individuals through the amalgamation of diverse data sets. This process could involve initially anonymized data sets containing personal information being combined with openly accessible data. This inadvertent exposure of personal details should always be considered a viable risk. As such, the precautionary measures adopted on both client and server sides should address this issue, and aim to prevent the reconstitution of any data or inference of the individuals to whom the data refers. Additionally, considering the high complexity of the operations to be conducted locally and models to be merged centrally, with the added difficulty to trace back the raw training data per the nature of federated data models, each training participant should abide by principle of accountability under Article 5(2) GDPR. This principle establishes the controller's responsibility to abide by principles relating to personal data and an ability to demonstrate continuing compliance to such principles, i.e., keeping records of the processing activities conducted on personal data especially containing "description of the categories of data subjects and of the categories of personal data"¹⁸⁵ to ensure a certain level of transparency. Thus, it would also be advisable to the TRUSTEE system developers, concerning the diligence required by the GDPR in practice, to be diligent with any record keeping possible to be able to explain the training processes behind the federated data model to be developed.¹⁸⁶ The partners are advised also to decide on a protocol demanding certain security measures to be put in place concerning the training data to be used and the model training processes to adhere to the requirement of maintaining a certain level of quality and accuracy of the data.

INTERNATIONAL DATA TRANSFERS AND SHARING AND CONSIDERATIONS ON DATA TRANSFER AND DATA SHARING AGREEMENTS

The aim for TRUSTEE solution in the future is to enable access to different types of data, provided by Data Providers to Consumers via TRUSTEE, regardless of their location.

In the first place, it is relevant to establish whether the data would be "transferred" with the support from TRUSTEE between the users, or whether making the data accessible to the Consumer at Data Provider's premises does not fulfil the requirement of "transfer", but is another type of "data sharing", and if so, then in what manner such data sharing should be governed.

From a technical perspective, "transfer" of the data would require a change of the geographical location of the data, whereas "data sharing" includes giving access to the data by whatever means.¹⁸⁷ The GDPR does not define data sharing itself, nevertheless, some DPAs consider it as "the disclosure of personal data by transmission, dissemination or otherwise making available"¹⁸⁸.

The GDPR addresses the transfer of personal data to a third country, but not "data sharing" with the entity based in the third country.

The relevant aspect which is of relevance is whether TRUSTEE allows for "data sharing" between separate controllers (Data Providers and Consumers) and if so, then what measures can be applied to appropriately govern this relationship. The code created by the UK ICO on data sharing is a helpful source of knowledge. The code indicates what processes are considered as data sharing. To these processes belong:

- *providing personal data to a third party, by whatever means.*

INFORMATICS, p.7. See also Jansi Rani Amalraj, Robert Lourdasamy, "Security and privacy issues in federated healthcare – An overview", Open Computer Science 2022 (12) 57–65, p. 64.

¹⁸⁵ GDPR, Article 30.

¹⁸⁶ Truong, Suna, Wang, Guittona, and Guo (n 4) 15.

¹⁸⁷ [Data sharing covered by the code | ICO](#).

¹⁸⁸ UK Data Protection Act 2018, Section 121.

- *receiving personal data as a joint participant in a data sharing arrangement.*
- *the two-way transmission of personal data; and*
- *providing a third party with access to personal data on or via controller's IT systems.*¹⁸⁹

Having considered the latter point, and the approach present in the TRUSTEE architecture, where TRUSTEE allows for accessing homomorphically encrypted personal data by Consumer on Data Provider's premises, it appears that TRUSTEE will facilitate "data sharing", and the parties using this solution will "share" their data between each other.

In such a situation, where sharing of the personal data would take place between two separate controllers, who would process the data for their own purposes, a "data sharing agreement" plays a relevant role. It is not directly defined in the GDPR (the GDPR defined only the standard contractual clauses related to "transfer" of personal data to third countries), nevertheless, its relevance is indicated by selected Data Protection Authorities (see UK ICO). TRUSTEE seem to assume this approach by providing a component which allows for conclusion of agreements between Data providers and Consumers for data access and use (STM).

According to the code issued by the UK ICO, a data sharing agreement between two separate controllers enables achieving a clear understanding about the roles of the parties, may set out the purpose of the data sharing, may indicate what will happen with the data at different stages of data sharing, as well as may set standards required by both parties. It may create a framework which will be a basis for justification that both parties recognised and addressed data protection principles.¹⁹⁰

A basic data sharing agreement shall include the major points such as:

- the parties identification and information on the participants of data sharing,
- the purpose of data sharing defined in a clear manner, i.e., description of the aims for which data are shared, why sharing is necessary to achieve those goals and what benefits to individuals and/or to society the parties hope to bring,
- what data are going to be shared
- lawful basis for data sharing (both parties need to clarify it for themselves – Data providers the basis for making data available, ad Consumers for accessing the data)
- information if there is any type of special categories of data that would be shared
- procedures allowing for achieving compliance with the rights of individuals (data subjects)
- governance arrangements between the parties, including:
 - advice on what datasets can be shared by the Data Providers and Consumers
 - ensuring accuracy of the data shared by the Data Providers,
 - standards for recording data
 - rules for retention and deletion of shared data and procedures
 - common technical and security solutions including procedures with dealing with a breach of the agreement
 - procedures for addressing access requests, complaints or queries,
 - procedures in the case of termination of the data sharing.¹⁹¹

The data sharing agreement must be reviewed on a regular basis.

In the case the future developments of TRUSTEE cause the situation in which the conclusion would be that the personal data are "transferred, below considerations would apply:

¹⁸⁹ [Data sharing covered by the code | ICO.](#)

¹⁹⁰ [Data sharing covered by the code | ICO.](#)

¹⁹¹ Ibid.

First, it would be crucial that Data Provider shares the information whether the data that would be made available with the support of TRUSTEE to Consumers are anonymised or pseudonymised. Provisions on data transfer to non-EU countries pursuant to the Chapter V GDPR do not apply to anonymous data. Nonetheless, it should not be understood as a permission to share anonymised data without limitations, as the risk of re-identification needs to be considered at all stages. It is then recommended that the access to the data via TRUSTEE should therefore be granted to the users who went through a strict authentication and authorisation process and only if the users may ensure that the data will be processed by them in line with data protection principles and will be kept confidential, where required.

If the confirmation that the data made available via TRUSTEE means that it is fully anonymised, the following points must be considered. Article 44 GDPR regulates that “any transfer of personal data, which are undergoing processing or are intended for processing after transfer to a third country” is allowed only if the controller and the processor who are supposed to receive the data comply with the requirements stemming from Chapter V GDPR.

If the Data Provider would transfer the data via TRUSTEE to a Consumer located outside of the EEA, the requirements described below must be addressed:

- a. If the country where the Consumer is located is deemed as ensuring an adequate level of protection (Article 45 GDPR). The list adequate countries is available in the Official Journal of the European Union and on its website.
- b. Making sure that before and during data transfer the Consumer has provided appropriate safeguards and on the condition that enforceable data subject rights and effective legal remedies for data subjects are available (Article 46 GDPR).

The appropriate safeguards may be provided in the form of:

- a legally binding and enforceable instrument between public authorities or bodies.
- binding corporate rules in accordance with Article 47 GDPR.
- standard contract clauses adopted by the Commission.
- standard contract clauses adopted by a supervisory authority and approved by the Commission.
- an approved code of conduct together with binding and enforceable commitments to apply safeguards; or
- an approved certification mechanism together with binding and enforceable commitments to apply safeguards.

In the context of appropriate security of the data transfer and sharing, the TRUSTEE architecture foresees to include security measures such as, in particular, HE, advanced access/authentication system, construction of a certification model, privacy impact assessments, federated learning approach

DATA PROTECTION IMPACT ASSESSMENT

In the case of the processing of personal data, Article 35 GDPR requires controllers to carry out a Data Protection Impact Assessment (‘DPIA’) concerning the processing activities likely to result in a high risk for the data subjects, such as human data analysis and sharing. The DPIA enables individuals to identify and minimise or mitigate data protection risks. This process is a structured one and involves several steps identified in Article 35(7) GDPR.

The TRUSTEE solution is foreseen as a solution which should allow the future users to process real data provided by the data provider and shared with the Consumer via TRUSTEE data space as a data connector. It needs to be considered that those data may present personal data. Therefore, in that case a DPIA would be recommended to be conducted by the end users, as the controllers of the data shared via TRUSTEE.

To enable that, the TRUSTEE solution will include a component developed in WP4, T 4.1, which will help the user (Data Provider or Consumer) to assess whether the data processing actions which are foreseen, involving the use of TRUSTEE, are posing a high risk to the data subject of the data uploaded onto TRUSTEE and whether the risk may be appropriately mitigated.

SECONDARY USE OF PERSONAL DATA

Secondary use of data means processing of data for purposes other than the ones initially indicated, for which the data were collected from the data subjects. Secondary use of data is not impossible, but requires adherence to the rules governing the re-purposing and re-use of personal data.

Processing of personal data for the purposes different than the initial ones, for which the data were collected is allowed only where the secondary processing is considered compatible with the primary purposes. In such cases no separate legal basis (i.e. other than the legal ground which the initial processing and purpose was based on) is necessary.

In principle, secondary data use is allowed if:

- a data subject (an individual whose data are intended to be further processed) expressed their consent for further processing;
- a Union or EU Member State law allows for it. It is the law which constitutes necessary and proportionate measure to safeguard, among others, national security, defence, public security, protection of the data subject or the rights and freedoms of others, general public interest, especially important economic or financial interest including public health and social security¹⁹²
- further processing is planned for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes –this type of purpose shall by default be considered compatible with the initial purposes.¹⁹³ However, in this case entity using previously collected data shall apply appropriate safeguards for the rights and freedoms of the individuals, such as appropriate technical and organisational measures in place to address data minimisation principle. One of the measures could be pseudonymisation. If the second purpose does not require processing of the data that allow for identification of the person – it should be conducted in that manner.¹⁹⁴

In other secondary use cases, compatibility with the primary purpose has to be proven by the controller. In assessing this compatibility, the following points have to be considered by the entity processing the data:

- Any link between the initial purposes for which personal data have been collected and the secondary purposes of the intended further processing – Are both of them related to each other? Would the secondary purpose be a logical next step in data processing?
- The context of personal data collection, especially relationship between the controller and the individual and reasonable expectations of a data subject – would further processing relate to already existing relationship or be beneficial for it?
- The nature of personal data – is it a special category? Will you use less intrusive method?
- The possible consequences of the further processing for individuals
- The existence of appropriate safeguards, which may include pseudonymisation and encryption of the data.¹⁹⁵

If compatibility test is not passed, and the research exemption is not applicable, it means that for any new purpose of processing of previously collected personal data, a new legal basis has to be found by the entity conducting such processing.

Recital 50 GDPR clearly provides that where a different controller processes the data at hand for other purposes (even if for research purposes), , processing by such controller would not be recognised as further processing but as processing for such controller's primary (research) purposes, which is subject to all applicable conditions, in

¹⁹² GDPR, Article 6(4) GDPR in conjunction with Article 23(1).

¹⁹³ GDPR, Article 5(1)(b).

¹⁹⁴ GDPR, Article 89(1) GDPR.

¹⁹⁵ GDPR, Article 6(4).

particular the establishment of a legal basis as well as, in the context of special categories of data, an exemption under Articles 6 and 9(2) GDPR, respectively.

Following that, Consumer of the data made available by a Consumer will be required to address requirements set forth in the GDPR like a separate, new controller, whereas a Data provider, before making the data available via TRUSTEE to the Consumer, can consider whether this data processing activity is compatible with the initial purpose, and can be based on the same legal basis, accordingly.

ETHICS REQUIREMENTS IN HEALTH DOMAIN

As seen above, there are several requirements stemming from the GDPR on the use of data, applicable also in the HEALTH domain. Although the TRUSTEE project will test its technical solutions on fictitious datasets, for an effective legal and ethical framework governing development and maintenance of TRUSTEE, ethical requirements, and obligations applicable to involvement of information concerning health and genetics of an individual must be considered and implemented. Ethical standards in medical research serve to “promote and ensure respect for all human subjects and protect their health and rights”¹⁹⁶. The project stipulates that one of its missions foreseen for HEALTH domain is “to assess the functional and non-functional specifications of this platform through a series of tests that mimic realistic data processing workflows while also investigate how other datasets can be utilised in the health domain.”¹⁹⁷ In order to complete this mission and “mimic realistic data processing workflows”, ethical and legal requirements must be abode by just as if real data were used in the Use Case.

One of the main principles in ethics for healthcare and health research is obtaining and respecting consent of individuals. When it comes to health data, legal requirements protect individuals but the trust in the use of the data needs to be secured and acceptable also from the ethics perspective¹⁹⁸. Some terms and obligations of the GDPR and ethical principles are similar, however the objectives differ, notably the consent requirement. Indeed, the consent provided in the GDPR as a legal basis refers to the processing of personal data whereas the ethical consent refers to the willingness to participate in the study or clinical trial. In order to distinguish the two, “consent” as an ethical requirement will be referred to as “ethical consent”. Unlike the GDPR, there is no explicit definition of “ethical consent”.¹⁹⁹ Obtaining ethical consent is part of the principle of patient autonomy and basic human rights²⁰⁰ of the individuals participating in the health research study, who should have an opportunity to express their willingness to take part in the research, and acknowledge the requirements and risks associated with their participation. An informed consent permits to “prevent the unauthorized usage of a participant’s own data in ways that are unknown, poorly understood, or in conflict with the values and commitments of the data donor.”²⁰¹ Despite the lack of a clear definition of consent, the Declaration of Helsinki provided by the World Medical Association lists the requirements regarding informed consent from an ethical perspective²⁰² - a consent must be both “informed” and “voluntary”²⁰³. The information provided to the participant shall include the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliation of the researcher, benefits and risks as well as the right to refuse and withdraw from the study. Additionally, the participant must be informed on the possibility for the data

¹⁹⁶ WMA Helsinki Declaration, §7

¹⁹⁷ TRUSTEE Grant Agreement, p.12.

¹⁹⁸ Kalkman, S., Mostert, M., Gerlinger, C. *et al.* Responsible data sharing in international health research: a systematic review of principles and norms. *BMC Med Ethics* **20**, 21 (2019).

¹⁹⁹ According to Wu and al. 2019, the Declaration of Helsinki on Ethical Principles for Medical Research involving Human Subjects does not provide the definition of ethical consent.

²⁰⁰ Horton (2002), p. 107

²⁰¹ Riso, B., Tupasela, A., Vears, D.F. *et al.* Ethical sharing of health data in online platforms – which values should be considered?. *Life Sci Soc Policy* **13**, 12 (2017).

²⁰² WMA, Declaration of Helsinki, §§25-32

²⁰³ WMA, Declaration of Helsinki, §25

generated based on his health information will be used for secondary use and under which conditions. Based on this information, the participant gives a “voluntary” consent, i.e., freely-given, without any dependent relationship with the practitioner and without any duress.²⁰⁴ Additionally, participant has to have the possibility to withdraw his or her consent and leave the study at any time. If a participant decides to withdraw the ethical consent, then his/her participation to the study ends with immediate effect, but it does not affect the legal processing of the personal data unless the participant asks for the personal data to be erased. In order to protect privacy of the participants, data collected from the study should be de-identified at least, anonymised when possible. When data is collected to later be shared, the said data must be de-identified to protect study participants and they must be aware that they can decline²⁰⁵.

In TRUSTEE, the Data Providers of the data from the health domain shall consider whether consent given by the participants allow them to allow the access and use of such data by the Consumers, and this applies regardless whether or not the data are pseudonymised or anonymised afterwards. When providing participants with the information, the secondary use of data must be emphasized by the Data Providers. In the context of the STM, the Data Provider could be required, along with the requirements provided to the Consumer, to also indicate whether ethical consent allows Data provider to enable accessing the data by Consumers.

Furthermore, sharing data for clinical research may be performed based on several grounds, notably scientific, ethical and economic²⁰⁶. The principles surrounding data sharing aim at “maximizing health benefits or social value”, which can be achieved by the quality and comprehensiveness of the data to be shared but also the scientific validity and social value²⁰⁷. Regarding ethics in data sharing, determining what are the values at stake and which one are deemed are a “component of successful mechanisms of data sharing”²⁰⁸. The Global Alliance for Genomics and Health relies on eight ethical principles for data sharing and user’s experience²⁰⁹:

- (i) Respect for the data sharing and privacy preferences of participants
- (ii) Transparency of governance and operations
- (iii) Accountability to best practices in technology, ethics, and public outreach
- (iv) Inclusivity by partnering and building trust among stakeholders
- (v) Collaboration to share data and information to advance human health
- (vi) Innovation in order to develop an ecosystem that accelerates progress
- (vii) Agility to ensure we act swiftly to benefit those suffering with disease
- (viii) Independence by structure and governance

These principles, while comprehensive, are too vague and need to take into account the particular issues that may arise with the sharing of data²¹⁰. In the case of TRUSTEE, as already mentioned, the scope of ethical informed consent is relevant, along with transparency of data sharing, which means ensuring information and accountability in sharing the data as well as making sure of their quality. Since, the data sharing will be performed with the use of AI, additional ethics considerations must be taken into consideration. These are reflected in the further sections.

²⁰⁴ WMA, Declaration of Helsinki, §26

²⁰⁵ Bauchner H, Golub RM, Fontanarosa PB. Data Sharing: An Ethical and Scientific Imperative. *JAMA*. 2016

²⁰⁶ Ohmann C, Banzi R, Canham S, *et al*, Sharing and reuse of individual participant data from clinical trials: principles and recommendations *BMJ Open* 2017, p.2

²⁰⁷ Kalkman, S., Mostert, M., Gerlinger, C. *et al*. Responsible data sharing in international health research: a systematic review of principles and norms. *BMC Med Ethics* 20, 21 (2019).

²⁰⁸ Riso B, Tupasela A, Vears DF, Felzmann H, Cockbain J, Loi M, Kongsholm NCH, Zullo S, Rakic V. Ethical sharing of health data in online platforms - which values should be considered? *Life Sci Soc Policy*. 2017 Aug 21

²⁰⁹ Riso, B., Tupasela, A., Vears, D.F. *et al*. Ethical sharing of health data in online platforms – which values should be considered?. *Life Sci Soc Policy* 13, 12 (2017).

²¹⁰ Riso, B., Tupasela, A., Vears, D.F. *et al*. Ethical sharing of health data in online platforms – which values should be considered?. *Life Sci Soc Policy* 13, 12 (2017).

The TRUSTEE Platform, should test and for its future exploitation implement solutions for the cross-discipline federation of data. Such data may not be only consisting of personal data but also very well of non-personal data from various sectors, such as from automotive and space datasets as example pilot cases will use for the trials in TRUSTEE. The Regulation (EU) 2018/1807 on the free flow of non-personal data²¹¹ complements the GDPR to provide a legal framework that ensures unrestricted movement of different types of data within the Union, by specifically covering the free flow of data other than personal data.²¹² One of the main points of emphasis of the Regulation is the assurance of trust in the security of cross-border data processing in order to reduce the preference of data localization practices “as a proxy for data security”.²¹³ In this line, it prohibits any data localization requirements, either mandating the processing of data in a particular Member State or prohibiting the processing in other Member States, to ensure the free movement of non-personal data within the Union, except for the circumstances where such practice can be justified on grounds of public security and is proportional.²¹⁴ The Regulation applies when the services of processing²¹⁵ of non-personal data are provided to users, who reside or have an establishment in the Union, regardless of the place of establishment of the service provider.²¹⁶ It does not apply to the services of processing of non-personal data taking place outside of the EU. The Regulation adopts a broad definition of data processing services encompassing both the IT systems used, being either outsourced or located on the premises of the user, and the data processing services of varying intensities.²¹⁷ Hence, TRUSTEE data space may be considered as a type of data processing service.

When the dataset contains both personal data and non-personal data, the GDPR will only apply to the personal data part of the dataset. However, when such types of data are inextricable from one another, this should not prejudice the applicability of the GDPR.²¹⁸ The Commission in its guidance states that in such a case, the data protection rights and obligations envisaged under the GDPR will fully apply to the whole mixed dataset even “when personal

²¹¹ In addition to the Regulation itself, on 29 May 2019, the European Commission issued its guidance, as envisaged by the Regulation itself, on the interaction of this Regulation and the GDPR, especially as regards data sets composed of both personal and non-personal data: COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL, Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union, 29 May 2019, COM(2019) 250 final (“EC Guidance on Regulation 2018/1807”).

²¹² EC Guidance on Regulation 2018/1807, p. 5-6: „The non-personal data can be categorised by origin as: Firstly; data which originally did not relate to an identified or identifiable natural person, such as data on weather conditions generated by sensors installed on wind turbines or data on maintenance needs for industrial machines. Secondly; data which were initially personal data, but were later made anonymous. The ‘anonymisation’ of personal data is different to pseudonymisation, as properly anonymised data cannot be attributed to a specific person, not even by use of additional data and are therefore non-personal data.” Recital 9 of the Regulation itself further exemplifies non-personal data as “the aggregate and anonymised datasets used for big data analytics, data on precision farming that can help to monitor and optimise the use of pesticides and water, or data on maintenance needs for industrial machines.”

²¹³ Regulation 2018/1807, Recital 33.

²¹⁴ Regulation 2018/1807, Recital 18 and Article 4(1).

²¹⁵ Processing as defined under Regulation 2018/1807, Article 3(2) and Regulation 2016/679, Article 4(2).

²¹⁶ Regulation 2018/1807. Recital 15 and Article 2(1).

²¹⁷ Recital 17.

²¹⁸ Article 2(2). Please also see EC Guidance on Regulation 2018/1807, p. 9-10: “if the non-personal data part and the personal data parts are ‘inextricably linked’, the data protection rights and obligations stemming from the General Data Protection Regulation fully apply to the whole mixed dataset, also when personal data represent only a small part of the dataset... The concept of ‘inextricably linked’ is not defined by either of the two Regulations. For practical purposes, it can refer to a situation whereby a dataset contains personal data as well as non-personal data and separating the two would either be impossible or considered by the controller to be economically inefficient or not technically feasible.”

data represent only a small part of the dataset.”²¹⁹ Though the Commission in its guidance does not exactly define what being “inextricably linked” means, it states that such a concept refers to cases “whereby a dataset contains personal data as well as non-personal data and separating the two would either be impossible or considered by the controller to be economically inefficient or not technically feasible.”²²⁰ Additionally, if with the advancement of technological methods, anonymized data can be reversed into personal data, such data shall be treated as personal data and the GDPR shall become applicable again.²²¹

Additionally, the Regulation underlines the importance of data porting as a contributing element to the free flow of data, whereby ideally, the data processing service providers should provide easy data portability by their professional users²²² when they switch to other service providers.²²³

As far as the TRUSTEE platform and this Regulation are concerned, the TRUSTEE data space would be considered a service provider as it aims to provide a data processing service within its data space.²²⁴

The primary point to be mindful of would be ensuring the protection of personal data in cases where a dataset containing personal data and a dataset containing non-personal data are unrecognizably linked. This would be particularly relevant for the exploitation phase of the TRUSTEE solutions by the users, who would consult to the solution to perform search queries across the federated data repositories to access and utilize data across different sectors and domains.

Additionally, depending on the services to be provided as an end-solution of the TRUSTEE project and to increase the trust of the future end users, the codes of conduct developed by SWIPO (the working group on cloud switching and porting data) could be adopted while evaluating the possible approaches to make sure that TRUSTEE data space allows for the option to switch between different cloud service providers and data porting to its professional users.

LEGAL FRAMEWORK FOR DATA GOVERNANCE AND ACCESS INCLUDED IN DATA GOVERNANCE ACT, DATA ACT (PROPOSAL) AND OPEN DATA DIRECTIVE

OPEN DATA DIRECTIVE

The Directive on open data and the re-use of public sector information (“Open Data Directive”), amending the Directive 2003/98/EC on the re-use of public sector information, aims to enable the benefiting to the fullest from “the potential of public sector information” in the internal market by addressing the still-existing differences in “rules and practices in the Member States relating to the exploitation of public sector information resources.”²²⁵ Re-use is defined by the Directive as “use by persons or legal entities of documents” held by either public sector bodies

²¹⁹ EC Guidance on Regulation 2018/1807, p. 9.

²²⁰ EC Guidance on Regulation 2018/1807, p. 10.

²²¹ Regulation 2018/1807, Recital 9.

²²² The Regulation, under Article 3(8), defines a professional user as “a natural or legal person, including a public authority or a body governed by public law, using or requesting a data processing service for purposes related to its trade, business, craft, profession or task.”

²²³ EC Guidance on Regulation 2018/1807, p. 18-19: “Article 6 of the Free Flow of Non-Personal Data Regulation does not provide for a right for professional users to port data, but has a self-regulatory approach, with voluntary codes of conduct for the industry... In practice, the portability under Article 6 of the Free Flow of Non-Personal Data Regulation concerns business-to-business interactions between a professional user (which may in cases that include processing of personal data qualify as ‘controller’ in accordance with the General Data Protection Regulation) and a service provider (similarly, to be qualified in some cases as ‘processor’).”

²²⁴ Regulation 2018/1807, Article 3(4).

²²⁵ Open Data Directive, Recital 15.

or public undertakings for “commercial or non-commercial purposes other than the initial purpose” of their tasks through which the documents²²⁶ were produced.²²⁷

The Directive, though not imposing a general obligation to allow re-use²²⁸, specifically addresses the promotion of re-use of existing documents held by public sector bodies and by some public undertakings as specified under the Directive²²⁹, as well as research data²³⁰ that are (i) from publicly funded research and, (ii) the researchers, research performing organizations or research funding organizations of which “have already made them publicly available through an institutional or subject-based repository.”²³¹

There are several exemptions to which the provisions do not apply, listed in Article 1 of the Directive.

Public sector bodies and public undertakings are expected to make their documents available, “where possible and appropriate, by electronic means, in formats that are open, machine-readable, findable and re-usable, together with their metadata.”²³² The principle to be followed in this regard is the principle of “open by design and by default.”²³³ As regards research data, the principle advised is “as open as possible, as close as necessary”, indicating that while Member States should adopt open access policies for publicly funded research data in accordance with FAIR principles, such policies should also address any concerns regarding IP rights, confidentiality, security and the protection of personal data.²³⁴ Public sector bodies are also obliged to make dynamic data available through appropriate technical means immediately after their collection. Dynamic data is defined as “documents in a digital form, subject to frequent or real-time updates, in particular because of their volatility or rapid obsolescence” (i.e., Data generated by sensors as an example given by the Directive).²³⁵

The re-use of documents are to be free of charge, though some public sector bodies and undertakings can benefit from an exception.²³⁶ Any marginal costs incurred from, for instance, anonymization of personal data, protection of commercially confidential information and dissemination of documents may be permitted.²³⁷ In addition, the re-use of documents should also not be subject to conditions, unless they are “objective, proportionate, non-

²²⁶ Documents, according to Article 2(6) of the Directive, can be any content regardless of their medium (i.e. paper or electronic form) or any part of their content.

²²⁷ Open Data Directive, Article 2(11).

²²⁸ Open Data Directive, Recital 26.

²²⁹ Open Data Directive, Article 1(1). Public undertakings specified by the Directive are: (i) the ones which are active in the areas defined in Directive 2014/25/EU on procurement by entities operating in the water, energy, transport and postal services sectors; (ii) the ones which are acting as public service operators pursuant to Article 2 of Regulation No 1370/2007 on public passenger transport services by rail and by road; (iii) the ones acting as air carriers fulfilling public service obligations pursuant to Article 16 of Regulation No 1008/2008 on common rules for the operation of air services in the Community and; (iv) the ones acting as Community shipowners fulfilling public service obligations pursuant to Article 4 of Regulation 3577/92 applying the principle of freedom to provide services to maritime transport within Member States.

²³⁰ Research data is defined under Article 2(9) as “documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results.”

²³¹ Open Data Directive, Article 10(2).

²³² Open Data Directive, Article 5(1).

²³³ Open Data Directive, Article 5(2).

²³⁴ Open Data Directive, Article 10(1).

²³⁵ Open Data Directive, Article 2(8).

²³⁶ Open Data Directive, Article 6.

²³⁷ Ibid.

discriminatory and justified on grounds of a public interest objective” as well as not being unnecessarily restrictive of the possibilities for re-use.²³⁸

The Directive also governs a special category of datasets as “high-value datasets” and lists in Annex I of the Directive a set of thematic categories of such datasets, which are geospatial, earth observation and environment, meteorological, statistics, companies and company ownership, and mobility data.²³⁹ Such datasets are defined as “documents the re-use of which is associated with important benefits to the society, the environment and the economy”, such as for instance the creation of value-added services.²⁴⁰ The Commission, according to Article 14(1), will soon adopt an implementing act regarding the high-value datasets under the categories listed above, which are to be followed by an impact assessment.

The Member State transpositions of this Directive should always be taken into account in the consideration of this Directive.

It should be further considered by the Partners involved in the development of TRUSTEE, based also on the feedback from the pilot use cases, whether and how the TRUSTEE solutions could incorporate any application or tool specifically addressing the facilitation of the re-use of data held by public sector bodies as envisaged by this Directive, as such bodies may potentially be the end users of the TRUSTEE data space.

DATA GOVERNANCE ACT

Data Governance Act²⁴¹ (“DGA”) regulates the conditions for the re-use of certain categories of public sector data and the framework for the conditions and monitoring of data intermediation services. The act applies to personal as well as non-personal data – “any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording”²⁴².

The DGA sets the framework of the rules on re-use of certain types of data held by public sector bodies. The scope of the data held by public sector bodies is not limited to only personal data but includes non-personal data as well. This part, however, plays a significant role in establishing the requirements for the end users, who may assume the roles of public sector bodies or data holders or data (re-)users. With respect to non-personal data, the DGA evaluates it from two different perspectives: on one part, the DGA underlines the importance of transmitting non-personal data only when the combination of various non-personal data sets would not lead to the identification of data subjects.²⁴³ On the other part, it highlights the significance of paying respect to the intellectual property rights or trade secrets residing in the non-personal data and of the contractual guarantees to be set to prevent any unlawful or unauthorized access to the non-personal data.²⁴⁴

More importantly for this project, the DGA further exemplifies data intermediation services²⁴⁵ as “data marketplaces on which undertakings could make data available to others, orchestrators of data sharing ecosystems that are open

²³⁸ Open Data Directive, Article 8.

²³⁹ Open Data Directive, Chapter 5 and Annex I.

²⁴⁰ Open Data Directive, Article 2(10).

²⁴¹ Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act).

²⁴² DGA, Article 2(1).

²⁴³ DGA Recital 15, Article 3(5).

²⁴⁴ DGA, Recital 20, Article 5(a)(ii), 5(3). Such guarantees can include, according to the DGA, modifying the data to prevent the disclosure of any confidential information, accessing the data within a remote premise and securing processing environment provided or controlled by the public sector body.

²⁴⁵ See definition in the Glossary.

to all interested parties, for instance in the context of common European data spaces”²⁴⁶. It however excludes from the definition the provision of services focusing on the “intermediation of copyright protected content”, “data sharing services that do not aim to establish commercial relationships”, services obtaining data to aggregate, enrich or transform the data in order to add substantial value to it and “license the use of the resulting data to data users, without establishing a commercial relationship between data holders and data users.” Thus, the DGA focuses its scope on data intermediation services that aim to assist commercial relationships between multiple parties, data holders, data subjects and data (re-)users. The DGA does not regard “cloud storage, analytics, data sharing software, web browsers, browser plug-ins or email services” as data intermediation services, “provided that such services only provide technical tools for data subjects or data holders to share data with others, but the provision of such tools neither aims to establish a commercial relationship between data holders and data users nor allows the data intermediation services provider to acquire information on the establishment of commercial relationships for the purposes of data sharing.”²⁴⁷ Having considered above details of the data intermediation services definition and further clarification, it may be considered that TRUSTEE could represent a type of such a service once it is put into service and would allow to establish commercial relationships between the data holders (in TRUSTEE represented by Data Providers) and data users (in TRUSTEE - Consumers). This assessment is based on the fact that, although TRUSTEE itself does not foresee to become a “data marketplace on which the data may be made available”, however, it is not excluded that TRUSTEE could be considered as “orchestrator of data sharing ecosystem”, by providing set of tools which could serve Data Providers and Consumers in making the data accessible in their relationship and ready for use.

The articles that are of relevance to the TRUSTEE solution are the ones setting the standards and framework for data intermediation services. The DGA subjects “intermediation services between data holders and potential data users, including making available the technical or other means to enable such services”, which may also be including “the creation of platforms or databases enabling the exchange or joint use of data, as well as the establishment of other specific infrastructure for the interconnection of data holders with data users”²⁴⁸ to a notification procedure. The data intermediation services are subject to a list of conditions as detailed above, such as having fair commercial terms, suitable formatting of data to be exchanged, fair and transparent procedures to access the service, interoperability with other data intermediation services, appropriate level of security, and providing the service in a neutral manner, to be able to provide this service.²⁴⁹

The DGA necessitates the data intermediation service providers who intend to provide the services listed above to submit a notification to the authority in the Member State where they have their main establishment.²⁵⁰ The provider of the data intermediation services, and their providers are subject to a set of obligations. According to those, the data sharing service providers are required to:

- safeguard trust and control over the data sharing between data holders and data consumers by ensuring the neutrality of the data intermediary.
- not use the data exchanged for any other purpose but only act as an intermediary;²⁵¹
- set the commercial terms for the provision of the service not dependent on whether the data holder or user uses any other services provided by the data intermediation service provider²⁵²;

²⁴⁶ DGA, Recital 28.

²⁴⁷ DGA Recital 28.

²⁴⁸ DGA Article 10(a).

²⁴⁹ DGA Article 12.

²⁵⁰ DGA Article 11.

²⁵¹ DGA Article 12(a), Recital 33.

²⁵² DGA Article 12(b).

- facilitate the exchange of the data in the format in which it receives it from a data subject or a data holder, and shall convert the data into specific formats only to enhance interoperability within and across sectors²⁵³;
- if wanted, offer additional services to data holders or data subjects that facilitate the exchange of data, such as temporary storage, curation, conversion, anonymisation and pseudonymisation; such tools should be used only at the explicit request or approval of the data holder or data subject and third-party tools offered in that context should not be used for other purposes²⁵⁴;
- apply suitable measures to ensure the interoperability of its services with other data intermediation services²⁵⁵;
- inform data holders, without delay, about an unauthorised access, transfer or use of the non-personal data that it has shared²⁵⁶;
- take necessary measures to ensure an appropriate level of security for the storage, processing and transmission of non-personal data²⁵⁷;
- further ensure the highest level of security for the storage and transmission of competitively sensitive information.²⁵⁸

The relevance of the DGA concerns more predominantly the aftermath of the project for the exploitation of the TRUSTEE solution, once TRUSTEE becomes exploited in a commercial manner. Nevertheless, in the development phase TRUSTEE system developers should consider these requirements which may have an impact on the technical development of tools planned to become a part of TRUSTEE architecture, to ensure that the system will be compliant with the DGA set of rules.

DATA ACT (PROPOSAL)

The Data Act's²⁵⁹ main aim is to provide and harmonise the rules on fair access to and use of data by consumers and businesses and to ensure fairness in the allocation of value from data among actors in the data economy, as well as to foster access to and use of data²⁶⁰ in the context of cloud services and IoT. It must be emphasised that at the current stage of the TRUSTEE development, the act itself is still undergoing legislative procedure, and the requirements stemming from the last update of the legal act (the version adopted by the European Parliament in March 2023) may still be subject to change. The changes will be reflected in the next versions of the Deliverable, nevertheless, initial framework as well as requirements stemming from the Act have been further analysed, as they may have impact on the development of the TRUSTEE architecture.

The term “data” is defined here in a similar manner as in the DGA – hence the goal is that the act will apply to personal and non-personal data. The difference in the EP's version of the Data Act in comparison to its first, proposal version is that it excludes “content or data obtained generated or collected by the connected product or transmitted to it on behalf of others for storage and processing purposes”.²⁶¹ This piece of legislation appear to be relevant for TRUSTEE and its users due to the fact, that it focuses on provision on harmonised rules on, in particular, data holders who make available data which they accessed from a product or generated during a provision of related

²⁵³ DGA Article 12(d).

²⁵⁴ DGA Article 12(e).

²⁵⁵ DGA Article 12(i).

²⁵⁶ DGA Article 12(k).

²⁵⁷ DGA Article 12(l).

²⁵⁸ DGA Article 12(l).

²⁵⁹ Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act) (Text with EEA relevance) {SEC(2022) 81 final} - {SWD(2022) 34 final} - {SWD(2022) 35 final}, COM(2022) 68 final 2022/0047 (COD), after amendments provided by the European Parliament, available at: [Texts adopted - Data Act - Tuesday, 14 March 2023 \(europa.eu\)](#)

²⁶⁰ [The Data act \(europa.eu\)](#), p. 3.

²⁶¹ Data Act, Article 2(1)

services to the users or to data recipients; fair contractual terms for data sharing agreements, which are of great relevance for the actors making use of TRUSTEE (Data Providers and Consumers); providing data as available to public entities in the case of an exceptional need in the public interest; the development of interoperability standards and specifications for data foreseen to be transferred and used.

The scope of subjects to whom the provisions defined in specific chapters apply, is broad:

- a) manufacturers of connected products and providers of related services on the EU market,
- b) users of such products and services as well as data subjects of personal data obtained, collected, or generated by the use of the product or service,
- c) data holders making data available to data recipients in the EU,
- d) data recipients to whom the data are made available,
- e) public sector bodies and other agencies and bodies which require data holders to make data available where an exceptional need in the public interest arises,
- f) providers of data processing services, regardless of where established, which offer such services to the customers within the EU.

To understand to what actors in TRUSTEE the Data Act applies, it is necessary, to better understand who stays behind the terms which Data Act introduces. The following subjects have to be further analysed: data holders, data recipients²⁶², users – all defined in relation to the connected products and related services and relevant for further analysis of the obligations of TRUSTEE users (i.e. Data providers and Consumers), and data processing services and their providers²⁶³, necessary to understand what obligations TRUSTEE system developers may have when developing and further maintaining and providing the infrastructure as a service.

The “data holder” (being a legal or natural person) is “an entity who has accessed data from the connected product or has generated data during the provision of a related service and who has the contractually agreed right to use such data, and the obligation, in accordance with the Data Act, , applicable Union law or national legislation implementing Union law to make available certain data to the user or a data recipient.”²⁶⁴ Considering the definition of the “connected product”²⁶⁵ and the “related service”, it seems to be, at least having considered the current shape of the Data Act, adopted by the European Parliament, as well as the characteristics of TRUSTEE and the entities who may use it for data access and (re-)use, based on the provided Use Cases described in Part II the Data Providers in TRUSTEE could potentially fall within the scope of this definition.

A “user” can be a legal or natural person. It owns a connected product or receives a related service or temporarily uses a product or receives service based on an agreement. A user can be also the data subject whose personal data are processed by a product or service.²⁶⁶

“Data recipient” is a person (legal or natural) to whom a data holder makes available data accessed from a connected product or generated during the provision of a related service. This person is someone else than a “user”. To be able

²⁶² Data Act, Article 2(7).

²⁶³ Data Act, Article 2(12).

²⁶⁴ Data Act, Article 2(6).

²⁶⁵ Under Article 2 of the Data Act Proposal, ‘connected product’ means ‘an item, that obtains, generates or collects, accessible data concerning its use or environment, and that is able to communicate data via an electronic communications service, a physical, connection or on-device access and whose primary function is not the storing, processing or transmission of data on behalf of others’; and ‘related service’ means ‘a digital service, including software , but excluding electronic communication services which is inter-connected with a product in such a way that its absence would prevent the product from performing one or more of its functions, and which involves accessing data from the connected product by the provider or the service’.

²⁶⁶ Data Act, Article 2(5).

to make data accessible to the data recipient, the data holder must obtain an explicit request by the user to do so, or this activity must be conducted due to a legal obligation defined either in the EU or national law²⁶⁷. Technically, considering the roles in TRUSTEE, a (TRUSTEE) Consumer could be recognised as a “data recipient” according to the Data Act, to whom the data would be made available by a data holder.

Further discussion on and performance of the scenarios for testing TRUSTEE, especially in the AUTOMOTIVE environment, will allow the project to better assess whether this initial consideration is accurate. The outcomes will be further presented in Deliverable 2.2 or 2.3 (presenting next version of the subject Deliverable).

What is crucial from the perspective of the TRUSTEE system developers and providers is to understand to what extent TRUSTEE could be recognised as a “data processing service”, defined as a digital service²⁶⁸, provided to a customer, which enables on-demand administration and broad remote access to a scalable and elastic pool of shareable computing resources of a centralised, distributed or highly distributed nature.²⁶⁹ The Data Act details that the mentioned access should be ubiquitous and on-demand, where “ubiquitous” means that “the computing capabilities are provided over the network and accessed through mechanisms which use heterogeneous thin/thick client platforms, going from web browsers to mobile devices and workstations. The catalogue of computing resources which are considered here is a broad one and includes solutions such as networks, servers, other virtual or physical infrastructure, software – and this includes software development tools, storage, applications and services. Scalability of the resources means that they should be flexibly allocated by the provider of services, regardless the geographical location of resources, to enable fluctuations handling in demand. “The term ‘shared pool’ is used to describe those computing resources that are provided to multiple users who share a common access to the service, but where the processing is carried out separately for each user, although the service is provided from the same electronic equipment.”²⁷⁰ Having considered the aim set out by TRUSTEE, it seems that the current draft definition could cover TRUSTEE. TRUSTEE aims to create a framework, consisting of set of elastic (and scalable) tools which would correspond with the “scalable computing resources provided in a distributed manner”, and provided to the TRUSTEE users for their further use according to their needs, and at the time when the user requests it, allowing to a) make data available by Data Providers and re-use them by Consumers, b) make AI models shared by Model Providers available as a service for further use and training by Consumers and c) further develop the tools and implement them in other, external infrastructures, in a manner that is the most suitable from the perspective of Developers.

Further term which plays a significant role in understanding what entity TRUSTEE and TRUSTEE system developers represent, is ‘common European data spaces’ introduced in Article 2 Data Act by the European Parliament. The current definition of the term reads ‘purpose- or sector-specific or cross-sectoral interoperable frameworks of common standards and practices to share or jointly process data for, inter alia, development of new products and services, scientific research or civil society initiatives.’²⁷¹ Considering a broad understanding of the term “data sharing”, presented in the section concerning “data sharing agreements” above, theoretically, TRUSTEE could be considered as a cross-sectoral interoperable framework of common standards and practices to share data for e.g. scientific research or development of new products or services, depending on the needs of TRUSTEE users. Therefore, the interoperability obligations posed for ‘operators of data spaces’ under the Data Act²⁷² are taken into account at the current stage of the project and further developments which would exclude their applicability will be reflected in Deliverable 2.1 and 2.3.

²⁶⁷ Data Act, Article 2(7).

²⁶⁸ Other than an online content service as defined in Article 2(5) of Regulation (EU) 2017/1128.

²⁶⁹ Data Act, Article 2(12).

²⁷⁰ Data Act, Recital 71.

²⁷¹ Data Act, Article 2(20a).

²⁷² [CiTiP White Paper Data Act \(kuleuven.be\)](https://kuleuven.be/citip/white-paper-data-act), p. 15, see Data Act, Article 28(1)

The vision of the data spaces is also available in the Commission Staff Working Document on data spaces. According to this document, a data space definition includes the entities offering “the deployment of data sharing tools and services for the pooling, processing and sharing of data by an open number of organisations, as well as the federation of energy-efficient and trustworthy cloud capacities and related services”, “data governance structures [...]”, and which “aim at improving the availability, quality and interoperability of data [...]”.²⁷³

There are important parts of this legislation act which may become relevant for TRUSTEE system developers and TRUSTEE users, once the Act is adopted in the proposed shape:

- (i) additional rules for B2B data sharing situations to make sure that data is shared under fair, reasonable, non-discriminatory and transparent terms²⁷⁴;
- (ii) rules on lock-in vendor issues in cloud computing.²⁷⁵ With the Data Act proposal, the EC engages much more decidedly with a detailed set of rules which effectively amount to a ‘right to switching’.²⁷⁶ In other words, the data processing services, such as, e.g. TRUSTEE, would be required to adhere to the conditions listed in Article 24, 25 and 26 of the Data Act and ensure that the technical, contractual and organisational obstacles for effective switching between the providers are removed.

The contracts between customers and providers of data processing services shall include:

- a) clauses allowing to switch to a data processing service offered by another provider or to port all data, applications and digital assets generated by the customer on an on-premise system;
- b) an exhaustive specification of all data and application categories exportable during the switching process, at minimum – all data imported by the customer and all data and metadata created by the customer and by the use of service (configuration parameters, security settings, access rights and logs);
- c) a minimum period of data retrieval of at least 30 calendar days.²⁷⁷

Additionally, providers of selected data processing services²⁷⁸ shall ensure that the customer, after switching to a service covering the same service type offered by a different provider of data processing services, enjoys functional equivalence in the use of the new service.²⁷⁹

All other providers shall make open interfaces publicly available and free of charge and ensure compatibility with open interoperability specifications of European standards for interoperability and if such standards do not exist, the data should be provided in a structured, commonly used and machine-readable format.²⁸⁰

- (iii) safeguards and more generally the rules on international data transfer by data processing service providers, to act against unlawful data transfer without notification by providers of data processing services. It

²⁷³ Commission Staff Working Document on Common European Data Spaces, SWD (2022) 45 final, 23.2.2022, Section 2.

²⁷⁴ Data Act, Chapter IV.

²⁷⁵ Data Act, Chapter VI.

²⁷⁶ [CiTiP White Paper Data Act \(kuleuven.be\)](https://kuleuven.be), p. 13.

²⁷⁷ Data Act, Article 25.

²⁷⁸ That means providers of services that concern scalable and elastic computing resources limited to infrastructural elements such as servers, networks and the virtual resources necessary for operating the infrastructure, but that do not provide access to the operating services, software and applications that are stored, otherwise processed, or deployed on those infrastructural elements.

²⁷⁹ Data Act, Article 26(1).

²⁸⁰ Data Act, Article 26 (2),(3),(4).

complements the legal regime of international transfers of non-personal as well as personal data and can thus be viewed as their continuation.²⁸¹

Article 27(1) Data Act requires the providers of data processing services to take all reasonable technical, legal and organisational measures, including contracts, to prevent international transfer or third-country governmental access to non-personal data held in the EU, where that transfer would cause a conflict with EU or national law. The paragraphs (2) and (3) provide additional requirements referring to the governmental access to the data.

- (iv) essential requirements (standards) to be complied with regarding interoperability of data spaces²⁸², addressed to participants of data spaces offering data or data services to other participants, to facilitate interoperability at the level of data, data sharing mechanisms and services (including the technical means to do so):²⁸³
- the dataset content, use restrictions, licences, data collection methodology, data quality and uncertainty shall be sufficiently described in a machine-readable format to allow the recipient to find, access and use the data²⁸⁴;
 - the data structures, data formats, vocabularies, classification schemes, taxonomies and code lists shall be described in a publicly available and consistent manner²⁸⁵;
 - the technical means to access the data, such as application programming interfaces, and their terms of use and quality of service shall be sufficiently described to enable automatic access and transmission of data between parties, including continuously or in real-time in a machine-readable format (where technically feasible and does not hamper good functioning of the product)²⁸⁶;
 - the means to enable the interoperability of contracts for data sharing within their services and activities shall be provided²⁸⁷.

These requirements may concern only specific sectors or, quite the opposite, have a generic nature and be appropriately interpreted, also taking into account other relevant sectoral legislation at the European and national level. The list is not closed, as it can be further complemented by the European Commission.²⁸⁸

The newest version of the act leaves some freedom with regards to the accountability rules concerning fulfilment of these requirements by the participants – the rules should be defined and agreed on by the participants within a particular data space.²⁸⁹

- (v) Article 30²⁹⁰, includes essential requirements set for smart contracts envisioned for data sharing. According to that, the entity offering smart contracts for others in the context of an agreement to make data available shall:

²⁸¹Data Act, Article 27, [CiTiP White Paper Data Act \(kuleuven.be\)](https://kuleuven.be/citip/white-paper-data-act), p. 14

²⁸² Data Act, Article 28.

²⁸³ Data Act, Article 28(1)..

²⁸⁴ Data Act, Article 28(1)(a).

²⁸⁵ Data Act, Article 28(1)(b).

²⁸⁶Data Act, Article 28(1)(c).

²⁸⁷ Data Act, Article 28(1)(d).

²⁸⁸ Data Act, Article 28(2).

²⁸⁹ Data Act, Article 3a.

²⁹⁰ The proposal of the European Commission included also the requirements: to foresee, if a smart contract must be terminated or deactivated, a possibility to archive transactional data, the smart contract logic and code to keep the record of the operations performed on the data in the past (**auditability**) (**data archiving and continuity**); to perform a conformity assessment with the view to fulfilling the requirements i-iv and issue an EU declaration of conformity – by doing this, it confirms its responsibility for compliance with these requirements.

- “ensure that the smart contract has been designed to offer a very high degree of robustness to avoid functional errors and to withstand manipulation by third parties“ (robustness)²⁹¹;
- “ensure that a smart contract is protected through and designed to offer rigorous access control mechanisms at the governance and smart contract layers” (access control)²⁹²;
- “ensure that a mechanism exists to terminate the continued execution of transactions: the smart contract shall include internal functions which can reset or instruct the contract to stop or interrupt the operation to avoid future (accidental) executions; in this regard, the conditions under which a smart contract could be reset or instructed to stop or interrupted, should be clearly and transparently defined. Especially, it should be assessed under which conditions non-consensual termination or interruption should be permissible” (safe termination and interruption)²⁹³;
- make sure that a smart contract affords “the same level of protection and certainty as any other contracts generated through different means” (equivalence)²⁹⁴
- “ensure that a smart contract respects and addresses the confidentiality of trade secrets” (protection of confidentiality of trade secrets)²⁹⁵.

Based on the current wording of the proposed Act, these requirements should be considered by TRUSTEE service developers/providers, taking into account that they will deploy smart contracts solutions, which TRUSTEE users will be offered to conclude agreements being the basis for making data available.

In the case a smart contract meets the harmonised standards drawn up and published in the Official Journal of the EU, it is presumed to be conform to the requirements.

Considering the support for interoperability and data openness, being an intrinsic part of the Data Act, reflected in Article 28 and 30 with regards to the interoperability obligations, as well as in Article 29 concerning open interoperability and portability specifications for data processing systems, there is no doubt that the Data Act supports data sovereignty and speaks against exclusive proprietary rights to data.

REGULATING AI – ANALYSIS OF INITIAL FRAMEWORK FOR AI SYSTEM PROVIDERS AND USERS

LAWFUL AI

AI processing should fulfil the already existing laws, to the great extent included in this Deliverable, including the primary law such as EU Treaties of the European Union and its Charter of Fundamental Rights, the secondary EU law, such as the General Data Protection Regulation, the Regulation on the Free Flow of Non-Personal Data, anti-discrimination Directives, consumer law, , the UN Human Rights treaties [with special attention for vulnerable groups such as children – the UN Convention on the Rights of the Child and the principles of the digital rights of the child], and the Council of Europe conventions (such as the European Convention on Human Rights), and various other EU Member State laws²⁹⁶. These are horizontally applicable laws.

²⁹¹ Data Act, Article 30(a).

²⁹² Data Act, Article 30(a).

²⁹³ Data Act, Article 30(b).

²⁹⁴ Data Act, Article 30(ba).

²⁹⁵ Data Act, Article 30 (bb).

²⁹⁶ To be further established with the support of pilot partners of the project. The main aspects relevant for TRUSTEE have been described in this legal framework analysis and will be further presented in the coming deliverables of the project.

TRUSTEE's objectives for data manipulation and AI processing will also have to adhere to several domain-specific rules that specifically apply in cases of AI use. For instance, the Medical Device Regulation²⁹⁷ in the healthcare sector or laws referring to AI processing in the transportation sector. While laws advance some parts of innovative practices, they also prohibit and hinder others. The challenges for TRUSTEE system developers but also TRUSTEE users are to have specific guidelines developed to inform TRUSTEE users and, where applicable, data subjects whose data TRUSTEE users are processing via TRUSTEE tools, if they are to deploy AI processing functionalities. The framework concerning AI application therefore is designed to ensure that any AI functionalities to be deployed at TRUSTEE will adhere to the right legal obligations.

AI system in TRUSTEE should also take into account approaching legislation, such as AI Act proposal, as well as relevant guidelines issued by authorities, concerning AI application and protection of data subjects.²⁹⁸

In the context of a lawful AI, i.e., compliant with applicable rules, in addition to the legal acts and regimes which are analysed in other sections of this legal framework, the subsection below elaborates on relevant aspects for provision and use of AI defined in the proposed AI Act. Further, the lawful AI considerations are complemented by the analysis how to achieve ethical and robust AI, and its core will be based on the Ethics Guidelines on Trustworthy AI issued by the European Commission.

Any further requirements relevant for AI from the legal perspective have been referenced in , and will be further complemented in the next stages of the project and discussed in future versions of this deliverable, namely D2.2 and D2.3.

AI ACT (PROPOSAL)

In April 2021, the European Commission published its Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence and amending certain union legislative acts (Artificial Intelligence Act) (hereinafter, "AIA"). Subsequently, next version of the AIA draft was presented in the Council General Approach document, published on 25 November 2022²⁹⁹, and including amendments adopted by the Council. The amendments of the European Parliament, which are planned to be voted upon in June 2023, are not included in the current shape of the following analysis, due to the fact that they have not yet been adopted by the whole plenary. They will be, however, taken into consideration after the amendments are adopted by the European Parliament, and updated considerations presented in Deliverable 2.2. The below analysis is therefore based on the AI Act version presented in the General Council Approach document, and whenever a reference "AI Act or AIA" appears, it refers to the version and the numbering of the Council amended version of this legal act.

The Act is to apply whenever an AI system is placed on the market or put into service within the EU, irrespective of where the provider of the system is based or where the system itself is based.³⁰⁰ If the users of the AI system are also located in the EU or the output produced by the AI system, the providers and users of which are located in a third country, is used in the EU, the Act still applies.³⁰¹

The Act defines the AI system as a system that:

²⁹⁷ Since the use case scenario does not refer to the use of the AI as described in the MDR, this will be included in the later iteration of D 2.1.

²⁹⁸ Selected guidelines, their scope and requirements stemming from them, will be further presented in D3.1.

²⁹⁹ The AIA in its newest version is available here: <https://artificialintelligenceact.eu/wp-content/uploads/2022/12/AIA-%E2%80%93-CZ-%E2%80%93-General-Approach-25-Nov-22.pdf>

³⁰⁰ AIA, Recital 11 and Article 2(1).

³⁰¹ AIA, Article 2(1).

“(i) is designed to operate with elements of autonomy; and

(ii) based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge-based approaches; and³⁰²

(iii) produces system-generated outputs such as content (generative AI systems), predictions, recommendations, or decisions, influencing the environments with which the AI system interacts”.³⁰³

The Council General Approach text³⁰⁴ includes an additional definition – “general purpose AI system” – which means an AI system that “is intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general purpose AI system may be used in a plurality of contexts and be integrated in a plurality of other AI systems”³⁰⁵, and all of it irrespective of how the system is placed on the market or put into service, including as open source software.

The provider of an AI system, according to the Act, is “a natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed and places it on the market or puts it into service under its own name or trademark, whether for payment or free of charge.”³⁰⁶

The Act proposes a risk-based approach, classifying the risks as unacceptable, high, limited, and minimal.³⁰⁷

An AI system would be considered as high-risk if the system is to be used as a product covered by the legislations listed under Annex II or to be used as a safety component of such a product, and if it is subject to a conformity

³⁰² This specific set of techniques used for the development of the software as the AI system were listed under Annex I of the AIA proposed by the Commission as: (a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; (c) Statistical approaches, Bayesian estimation, search and optimization methods. The European Commission has the authority to make any amendments to this list of techniques. However, the Council General Approach (see footnote 170) deleted Annex I.

³⁰³ This is the definition that was revised by the Council in its “General Approach”, published 25 November 2022 (“Council General Approach”). Before the General Approach was presented, the definition of the AI systems under Article 3(1) read as “a software that is developed with a specific set of techniques and can, “for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with”.

³⁰⁴ The definition is included in the text published on 25 November 2022, Article 3(1b). Text available at: <https://artificialintelligenceact.eu/wp-content/uploads/2022/12/AIA-%E2%80%93-CZ-%E2%80%93-General-Approach-25-Nov-22.pdf>.

³⁰⁵ Ibid.

³⁰⁶ AIA, Article 3(2).

³⁰⁷ European Parliament (2022). Legislative Train Schedule. Proposal for a regulation on a European approach for Artificial Intelligence in “A Europe Fit for the Digital Age”, available [online] at: <https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence>. The Commission proposes a limited set of obligations, such as transparency obligation, for limited risk AI systems and foresees no additional obligations than the existing ones under current legislation for minimal risk AI systems. Article 69 of the AIA envisages the drawing up of codes of conduct addressing AI systems other than high-risk AI systems to encourage the producers of such systems to voluntarily abide by the mandatory requirements set for high-risk AI systems.

assessment conducted by a third-party.³⁰⁸ The AI systems with high risks (also general-purpose AI systems being part of high risk AI) are subject to a set of requirements including:

- Establishing, implementing, documenting, and maintaining a continuous and iterative risk management system for the entire lifecycle of the high-risk AI system. Risk mitigation measures should cover the identification of known and foreseeable risks, elimination or reduction of risks through adequate design and development, mitigation and control measures for risks that cannot be eliminated, providing adequate information on the risks as well as training to users, if suitable.³⁰⁹ Residual risks must be communicated to users.³¹⁰ These systems must also be tested to detect most appropriate risk management measures.³¹¹ The risk management should also consider if the users are under age of 18.³¹²
- Ensuring that the training³¹³, validation³¹⁴ and testing³¹⁵ data sets meet a certain quality criteria, if the high-risk AI system uses techniques including the training of models with data³¹⁶. These data sets are required to be subject to data governance and management practices regarding, for instance, relevant design choices, data collection, data preparation processing operations³¹⁷, detection of any data gaps or shortcomings³¹⁸, examination in view of possible biases likely affecting health and safety of individuals and leading to discrimination prohibited by EU law³¹⁹ and a prior assessment of the availability, quality and suitability of data sets³²⁰. In addition, “[t]raining, validation and testing data sets shall be relevant, representative, free of errors and complete”.³²¹ The developer needs to adopt appropriate data governance and management practices in training, validating and testing data sets.
- Providing a technical documentation of the high-risk AI system before it is placed on the market or put into service and regularly updating such documents. The documentation needs to show how the high-risk AI

³⁰⁸ AIA, Article 6(1) and (2). The AI systems under Annex III of the AIA, however, are automatically considered as high-risk. Some examples include AI systems intended to be used for the real time and post remote identification of natural persons, systems used to assess students in educational and vocational training institutions, or systems used for recruitment or selection of natural persons for job vacancies.

³⁰⁹ AIA, Article 9.

³¹⁰ AIA, Article 9(4).

³¹¹ AIA, Article 9(5).

³¹² AIA, Article 9(8).

³¹³ Training data, under Article 3(29) of the AIA version included in the Council General Approach text, is defined as “data used for training an AI system through fitting its learnable parameters.”

³¹⁴ Validation data is, according to Article 3(30), any data “used for providing an evaluation of the trained AI system and for tuning its non-learnable parameters and its learning process.” Such data can belong to a separate dataset or be part of the training data set.

³¹⁵ Testing data, under Article 3(31), is described as “data used for providing and independent evaluation of the trained and validated AI system in order to confirm the expected performance of that system.”

³¹⁶ Quality criteria are defined in Article 10(2)-(5).

³¹⁷ AIA, Article 10(2)(c): “Such as annotation, labelling, cleaning, enrichment and aggregation”.

³¹⁸ AIA, Article 10(2)(g).

³¹⁹ AIA, Article 10(2)(f). According to Article 10(5), „To the extent that it is strictly necessary for the purposes of ensuring bias monitoring, detection and correction in relation to the high-risk AI systems, the providers of such systems may process special categories of personal data referred to in Article 9(1) of Regulation (EU) 2016/679, Article 10 of Directive (EU) 2016/680 and Article 10(1) of Regulation (EU) 2018/1725, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons, including technical limitations on the re-use and use of state-of-the-art security and privacy-preserving measures, such as pseudonymisation, or encryption where anonymisation may significantly affect the purpose pursued”. This is, however, still just a proposal.

³²⁰ AIA, Article 10(2)(e).

³²¹ AIA, Article 10(3).

system complies with the requirements set in AIA and set by any other competent authorities. Annex IV to the proposed AIA details what specific information the technical documentation must include.³²²

- Embedding in the AI system the ability to automatically record events (keep logs) as the system operates, enabling the “traceability of the AI system’s functioning” and also the “monitoring of the operation of the high-risk AI system with respect to the occurrence of situations that may result in the AI system presenting a risk.”³²³
- Designing and developing the AI systems in a proper manner to ensure that “their operation is sufficiently transparent to enable users to interpret the system’s output and use it appropriately.” Instructions on how to use these systems should also follow and be presented in an appropriate digital format. Next to these, information such as the high-risk AI system’s purpose, capabilities and limitations need to be communicated to its users in a concise, complete, correct and clear manner.³²⁴
- Implementing mechanisms to assure effective human oversight over the use of the AI system. The Act requires the assurance of this through the building of an oversight mechanism into the system by the provider before it is placed on the market or put into service, and/or if the former is not technical feasible, through measures that are appropriate to be implemented by the user of the AI system. The role of human oversight should give the individuals responsible of such oversight the capacity to, for instance, understand the limitations of the system, monitor the system’s operation to spot any dysfunctions, correctly interpret the outputs of the system, decide to opt out of using the system and its outputs in a specific circumstance, and to intervene on its operation through a “stop” button or any other measure alike.³²⁵
- Designing and developing the systems in a way to ensure an appropriate level of accuracy, robustness and cybersecurity. Robustness and resilience signify the necessity for the AI system to be adaptable against any errors, faults and inconsistencies and to withstand any attempts by unauthorized parties to alter the use or performance of the system.³²⁶
- Establishing and documenting a post-monitoring plan for high-risk AI systems to collect, document and analyse data provided by users or other data collected on the performance of the AI system throughout their lifecycle actively and systematically. This plan should be part of the technical documentation required from the providers of such systems.³²⁷

The providers of high-risk AI systems are also subject to a list of obligations themselves, including the application of a quality management system, implementing a conformity assessment procedure³²⁸ for the system before placing it on the market, and performing the necessary corrective actions if the AI system does not comply with the requirements under AIA.³²⁹

³²² AIA, Article 11 and Annex IV. In Annex IV, some of required pieces of information are descriptions on AI system’s intended purpose, its developers, the processes required for its development, the logic behind the design specifications and algorithms choices, and detailed information on the monitoring, functioning and control of the AI system.

³²³ AIA, Article 12.

³²⁴ AIA, Article 13. Article 52 further elaborates the transparency obligation for certain high-risk AI systems.

³²⁵ AIA, Article 14. Additionally, the user of an AI system is defined as “any natural or legal person, public authority, agency or other body using an AI system under its authority” under Article 3(4).

³²⁶ AIA, Article 15.

³²⁷ AIA, Article 61.

³²⁸ AIA, Article 17.

³²⁹ AIA, Article 16.

In addition, the AIA brings forth transparency obligations AI systems aiming to interact with natural persons, in order to clearly communicate to the users that they are interacting with an AI system, unless this is evident from the context in which the AI system is placed for use.³³⁰

The AIA also foresees the establishment of “regulatory sandboxes” either by the competent authorities of some Member States or by the European Data Protection Supervisor in order to “provide a controlled environment that facilitates the development, testing and validation of innovative AI systems” before they are placed on the market or put into service.³³¹

It should be noted here that, though not being present in the initial text, the Council General Approach text from 25 November 2022 introduced an exception for AI systems “developed and put into service for the sole purpose of scientific research and development”³³², akin to the logic of the exception offered in the GDPR under Article 89.

The mentioned obligations have been reflected in the current version of the legal and socio-ethical requirements included in Part II of the subject Deliverable. It is not excluded that TRUSTEE service providers could be potentially considered as AI providers or AI distributors; therefore, it is relevant that, along with the Model providers who will be sharing the AI models via TRUSTEE, also TRUSTEE system developers/providers should consider and apply approaches compliant with the provisions included in the AIA.

As the discussions on the proposal for the Artificial Intelligence Act are still ongoing between within the European Parliament and the Council, the text of the AIA and the rules within are still open to change.

SOCIO-ETHICAL, ROBUST AND SAFE AI

Data manipulation and AI processing will not only fulfil the obligations set by the law but also adhere to an ethical framework to mitigate any issues emerging in cases where laws are not up to speed with the technological developments and therefore become inadequate in certain aspects (or worse – step over certain ethical requirements). Therefore, for TRUSTEE to be trustworthy, the system must also align with ethical norms.

The TRUSTEE system should perform in a safe, secure, and trustworthy way, and adopt a risk-based approach to development. This means that an iterative methodology of identifying, analysing, and avoiding foreseeable and emerging risks should be adopted. It is therefore important to ensure that any AI processing deployed for the TRUSTEE is robust. TRUSTEE should demonstrate technical robustness as appropriate in specific contexts, such as the application domain or life cycle phase, and from a social perspective, the societal domain within which or for which TRUSTEE will operate.

These aspects are also addressed by the Ethics framework presented in Ethics Guidelines for Trustworthy AI³³³, further discussed in the section below, and being an orientation point for TRUSTEE system developers, as well as TRUSTEE users.

ETHICS GUIDELINES FOR TRUSTWORTHY AI

³³⁰ AIA, Article 52.

³³¹ AIA, Article 53.

³³² AIA, Recital 12b, and Article 2(6)(7). For a criticism on the still unclear and narrow wording of this exception, please see Kazim et al., Proposed EU AI Act—Presidency compromise text: select overview and comment on the changes to the proposed regulation, AI and Ethics (2022), accessed from <https://link.springer.com/article/10.1007/s43681-022-00179-z>.

³³³ <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

A key role for assessment of the requirements for AI, their providers and users play the Ethics Guidelines for Trustworthy AI, presented by the High-level Expert Group on AI (HLEG AI) (further referred to as “Guidelines”), in April 2019³³⁴. Following the Guidelines, AI which aims to be acknowledged as trustworthy shall be lawful (see sections above), ethical and robust.

The Guidelines do not address in detail the first component but focus more on the guidance on securing two latter components. The seven basic requirements, which, according to the Guidelines, should be taken into account and met when developing and deploying AI, which are also of utmost importance for TRUSTEE system developers and users. This list is further detailed by the Assessment List for Trustworthy Artificial Intelligence, i.e., a list consisting of points built on the requirements indicated in Guidelines for self-assessment targeting developers of and parties deploying artificial intelligence (“AI”).³³⁵

The requirements are as follows:

- Human Agency and Oversight - this requirement underlines the necessity for the AI system to have “the principle of respect for human autonomy” as a core principle, applied to its design. It refers to the effect of the AI system on human behavior, judgement, and decision-making, as well as trust and (in)dependence.³³⁶ The requirement necessitates putting in place mechanisms to prevent overreliance by users on AI systems and undesirable interferences by the AI systems on users’ decision-making. This sub-section also recommends the evaluation of certain governance mechanisms, such as “human-in-the-loop”, “human-on-the-loop” or “human-in-command”, through which human autonomy and ability to oversee and intervene in the AI system can be safeguarded.³³⁷
- Technical Robustness and Safety – it necessitates a risk-based approach during the development of the AI system, and an endeavour for the minimisation and prevention of any unintentional and unexpected harm. Goals such as the following should be aimed for during the development and deployment of the AI system:
 - o increasing the accuracy of the AI system by using data (including the training data used) that is “up-to-date, high quality, complete and representative of the environment the system will be deployed in”³³⁸ and by applying functions in which AI system can indicate how likely the inaccurate predictions are, and support, correct and mitigate risks from such predictions³³⁹;
 - o setting up mechanisms to ensure the security and resilience of the AI system and to protect it against any (cyber)attacks.

³³⁴ <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

³³⁵ <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

³³⁶ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 15 and 16; The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 7.

³³⁷ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 16 and The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 8: “Human-in-the-loop refers to the capability for human intervention in every decision cycle of the system. Human-on-the-loop refers to the capability for human intervention during the design cycle of the system and monitoring the system’s operation. Human-in-command refers to the capability to oversee the overall activity of the AI system (including its broader economic, societal, legal and ethical impact) and the ability to decide when and how to use the AI system in any particular situation.”

³³⁸ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 10. Please also see GDPR Article 5(1)(d).

³³⁹ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 17.

- applying safeguards enabling a fallback plan, inter alia, defining the risks and risk levels clearly³⁴⁰;
 - reducing the AI bias.
 - making the system reliable and resilient towards changes, as well as reproducible;³⁴¹
- Privacy and Data Governance - it is based on the fundamental right to privacy and prevention of any harm to it. Upon this basis comes the application of data governance regimes that this requirement necessitates, which should encompass analysing “the quality and integrity of the data used”³⁴² as well as of the “capability to process data in a manner that protects privacy”.³⁴³ In addition, ensuring appropriate governance of access rights to plays here a great role.³⁴⁴ One important checkbox that the target group should be aware of is whether, depending on the use case, “mechanisms that allow flagging issues related to privacy concerning the AI system” were established.³⁴⁵ Under the self-assessment list for the data governance framework, target group are expected to:
 - a) detect whether any use or processing of personal data took place while training or developing the AI system.
 - b) put in place any measures which are mandated or recommended by the GDPR, such as the Data Protection Impact Assessment (“DPIA”), oversight mechanisms for processing data (etc. “limiting access to qualified personnel, mechanisms for logging data access”), anonymization of data, pseudonymisation of data, “adequate, relevant and limited” processing of personal data (data minimisation)³⁴⁶, assigning of a data protection officer³⁴⁷;
 - c) “implement the right to withdraw consent, the right to object and the right to be forgotten into the development of the AI system”;
 - d) “consider the privacy and data protection implications of data collected, generated or processed over the course of the AI system’s life cycle”, not just the processing of personal data but non-personal data as well.³⁴⁸;
 - Transparency - this requirement includes three sub-elements, which are “traceability, explainability and open communication about the limitations of the AI system.”³⁴⁹ The first element requires clear and organized documentation of the processes applied while developing the AI system in order to establish trust in and transparency of the system. By documentation, the list recommends measures that will allow tracing back to, for instance, which data was fed into the AI system and which rules

³⁴⁰ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 16.

³⁴¹ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 16 and 17.

³⁴² Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 17.

³⁴³ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 12.

³⁴⁴ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 17.

³⁴⁵ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 12.

³⁴⁶ GDPR Article 5(1)(c).

³⁴⁷ GDPR Article 37.

³⁴⁸ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 13.

³⁴⁹ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 14.

were coded for the decision-making process behind the system.³⁵⁰ The second element, explainability, refers to the necessity to help users of the AI system understand why certain outputs and decisions are presented to them in order to allow them to contest any decision.³⁵¹ Third sub-element, communication, calls for the circulation of the information on limitations and risks, as well as benefits users should expect to face when interacting with the AI system.³⁵²

- Diversity, Non-discrimination and Fairness - This requirement demands a focus on the avoidance, minimisation and removal of the bias of the AI system concerning both “the use of input data as well as for the algorithm design” both during its training and operation.³⁵³ The expert group advises putting processes in place “to test and monitor for potential biases during the entire lifecycle of the AI system” and to “consider diversity and representativeness of end-users or subjects in the data.”³⁵⁴ It also advises putting in place “a mechanism to include the participation of the widest range of possible stakeholders in the AI system’s design and development.”³⁵⁵
- Environmental and Societal Well-Being – this set of requirements encourages AI systems to address sustainability and environmental friendliness aspects, and to consider and assess the positive and negative effects and impact the developed, deployed and/or used AI systems may have on democracy, institutions and society at large.³⁵⁶
- Accountability - this requirement entails assuring “responsibility for the development, deployment and/or use of AI systems”³⁵⁷, implementing risk management, independent/third-party audit and redress mechanisms and possibilities, especially incorporating these in the design of the AI system, minimising and reporting of negative impacts rational, as well as addressing of trade-offs of the requirements in the implementation phase.³⁵⁸

The list of the requirements is a flexible one, meaning that the target group can focus on the requirements especially relevant to their AI systems or even add any other requirements important to the sector in which they operate. The application of these requirements to the actions and solutions of the AI developers are seen as necessary to earn the trust of users and parties affected by the AI systems, the “trust that their (AI systems’) design, development and use

³⁵⁰ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 18.

³⁵¹ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 18.

³⁵² Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 18.

³⁵³ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 16; Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 18.

³⁵⁴ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 16; Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 18-19.

³⁵⁵ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 18; Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 19.

³⁵⁶ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 19.

³⁵⁷ The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p. 21.

³⁵⁸ Independent High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, 18 April 2019, p. 20.

are lawful (complying with all applicable laws and regulations), ethical (ensuring adherence to ethical principles and values), and robust.”³⁵⁹

Within its core ecosystem, TRUSTEE aims to entail a list of applications and services, one of these being the “Artificial Intelligence Infrastructure as a Service (AIaaS)”.³⁶⁰ The service will consist of multiple dimensions, such as “Explainable AI”, “Trustworthy AI”, and “Resilient AI”.

The service intends to enable researchers to create machine-learning workflows to test their own algorithms or use algorithms provided by TRUSTEE ecosystem on TRUSTEE datasets.³⁶¹ The project defines this service as a “trustworthy AI playground, where researchers can securely design, deploy, and test transparent AI-related models, algorithms, and architectures” and can model “eXplainable AI (xAI) to enforce transparency” and implement “privacy-friendly AI workflows”.³⁶² The Explainable AI tool is another model TRUSTEE aims as an output in order to assist providing trustworthiness and explainability of AI/ML models while the users perform big data analytics and extract knowledge from the data spaces.³⁶³ Finally, under its compliance and certification scheme, TRUSTEE aims to deliver a compliance support tool (CST – TAI-SDF) and provide an assessment of ethical AI by drafting a checklist in line with the Ethics Guidelines for Trustworthy AI and other applicable standards identified as relevant for TRUSTEE.³⁶⁴

This shows the inextricability of the TRUSTEE’s core objectives with the ALTAI requirements for a trustworthy and ethical AI. All the requirements listed above, especially privacy, data governance and transparency, should feed into not just the development of the AI services aimed as outputs of the project, but also to any type of action taken on data, particularly on personal data, throughout the project. Specifically speaking for the explainable AI tool, the extent and application of this service should be clearly outlined, and the explainability aimed should encompass both the technical process behind the AI system as well as the rationale behind the output the AI system gives.

ETHICAL ASPECTS OF XAI

In the modern era, AI systems can perform a large variety of tasks that require human intelligence and can bring tremendous changes in today’s society. Semi-autonomous or even autonomous AI based systems are employed in a large variety of domains, and in some cases their decisions can be of critical importance, e.g. Healthcare, transportation. As AI gains ground in applications that affect society, fears that its operation can have negative implications on human employment and interaction have emerged. For this reason, many organizations have developed expert committees on AI that have developed guidelines related to the ethical use of AI with the aim of addressing these fears. Such AI related ethical challenges concern both public and private entities³⁶⁵.

³⁵⁹ The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment, Independent High-Level Expert Group on Artificial Intelligence, 17 July 2020, p.4, referring to Ethics Guidelines for Trustworthy AI, Independent High-Level Expert Group on Artificial Intelligence, 8 April 2019.

³⁶⁰ TRUSTEE Grant Agreement Part B, p. 10, 14.

³⁶¹ TRUSTEE Grant Agreement Part B, p. 14.

³⁶² TRUSTEE Grant Agreement Part B, p. 14.

³⁶³ TRUSTEE Grant Agreement Part B, p. 14, 15.

³⁶⁴ TRUSTEE Grant Agreement Part B, p. 15. Note: The Assessment List For Trustworthy Artificial Intelligence (ALTAI) for self-assessment builds on the Ethics Guidelines for Trustworthy AI. The checklist/ TAI-SDF self-assessment steps and the basis for creation of such will be further reflected in D3.1.

³⁶⁵ Jobin, A., Ienca, M. & Vayena, E. The global landscape of AI ethics guidelines. *Nat Mach Intell* 1, 389–399 (2019). <https://doi.org/10.1038/s42256-019-0088-2>

The ethical requirements for an AI based system are requirements that derive from ethical principles. According to the European Commission's ethical guidelines for trustworthy AI³⁶⁶, these principles include: *Autonomy*, which is the respect for human dignity, *Beneficence*, which means acting with the aim of doing good to others, *Nonmaleficence*, which means not intentionally harming others, *Justice*, which means treating everyone fairly, and *Explicability*, which means having a transparent behaviour. By default, AI systems cannot act and take decisions based on ethical principles.

However, they have the functionality to meet certain ethical requirements that have been specified during their design, like the other requirements. By the definition of the ethical requirements, ethical issues are considered from the beginning of the development process. So, developers of AI systems and stakeholders that are involved should include these ethical issues during the specification of the requirements of the system. In order to derive the ethical requirements, it is important to take into consideration all the stakeholders that are involved in the social setting that the AI system is deployed.

These could be described as essential requirements that are derived from the whole ecosystem in which the AI system is applied. In addition, during the requirements evaluation phase, a focus on the ethical aspects is guaranteed. For the implementation of the functional and quality requirements that derive from the ethical principles, it should be noted that the AI system should behave like a properly trained human for this specific task. This means that the AI system should not only provide results, but it should also provide an explanation about how it came up with a particular prediction or decision, like a human operator would do³⁶⁷. This is where XAI techniques can contribute towards fulfilling ethical requirements.

Nowadays, when designing an AI model, experts usually face a trade-off between accuracy and interpretability. In most cases, simpler models are easier to understand but have a poorer performance, compared to complicated models that showcase high performance but are hard to interpret. Usually, accuracy and interpretability are not possible to be both maximized, and there is no universally accepted criterion by which of the two variables is more important. So, this trade-off in AI models imposes a not only technical, but also ethical, issue³⁶⁸. Taking this aspect into consideration, and in the context of TRUSTEE, the XAI tool shall be able to provide human understandable explanations for the end users, without sacrificing human interpretability in the favour of higher performance.

FROM FUNDAMENTAL RIGHTS TO SOCIO-ETHICAL PRINCIPLES FOR AI USE

The proposed framework that will guide both technical and non-technical stakeholders of TRUSTEE is based on the fundamental rights enshrined in the EU Treaties, the EU Charter, and the international human rights law. Ensuring that any data manipulation and AI functions are deployed on TRUSTEE have respect for fundamental human rights, within a framework underpinned by the rule of law, builds a solid foundation for further introducing and enforcing more abstract ethical principles and values, which will guide specifically the development and deployment of AI. As legally enforceable rights, the fundamental rights fall under the first component of this legal and ethical framework for data manipulation and AI use, making data manipulation and AI use lawful, which safeguards compliance with the law. Together with the HRESIA assessment and evaluation of data manipulation and AI use, provide an ethical component to the framework, therefore making data manipulation and AI use ethical.

³⁶⁶ "High-Level Expert Group on Artificial Intelligence, E.C.: Draft ethics guidelines for trustworthy AI", draft document (2018)

³⁶⁷ Guizzardi, R., Amaral, G., Guizzardi, G., Mylopoulos, J. (2020). Ethical Requirements for AI Systems. In: Goutte, C., Zhu, X. (eds) *Advances in Artificial Intelligence. Canadian AI 2020. Lecture Notes in Computer Science()*, vol. 12109. Springer, Cham. https://doi.org/10.1007/978-3-030-47358-7_24

³⁶⁸ Izumo, T., Weng, YH. Coarse ethics: how to ethically assess explainable artificial intelligence. *AI Ethics* 2, 449–461 (2022). <https://doi.org/10.1007/s43681-021-00091-y>

This second aspect of the proposed framework (and toolset especially for the technical stakeholders of TRUSTEE) deal with ethical aspects of data manipulation and AI which are not necessarily legally binding yet still incremental to ensure trustworthiness of the whole platform of TRUSTEE and its intended AI capabilities.

Nevertheless, even after complying with legally enforced fundamental human rights, ethical consideration for data manipulation and AI is crucial at this stage and as part of an evaluative and guiding framework or toolset for at least two reasons. First, a socio-ethical dimension will help both technical and non-technical stakeholders of TRUSTEE understand how the deployment of such functionalities as AI and data manipulation can implicate fundamental rights and their underlying values. And second, it can also help them provide more granular guidelines about what should be done and exist (as AI and algorithmic automation and data manipulation) rather than what technologies are presently capable of doing. In other words, the respect for human dignity and humans' best interests, freedoms, respect for justice and the rule of law, equality, non-discrimination and solidarity, and citizens' rights should be the priority when engineering AI functionalities. Certainly, while ethics can be seen as a 'soft' approach to regulation – a form of self-regulation whereby entities and projects (including like the TRUSTEE platform) can merely tick off the box as 'ethical' – socio-ethical guidance should not be left as the only one.

HUMAN RIGHTS IMPACT ASSESSMENT OF AI

While it is crucial to adhere human rights-oriented approach in data manipulation and AI functionalities for TRUSTEE, and the role that impact assessment processes can play in that regard, it should be noted that traditional Human Rights Impact Assessment (HRIA) models are territory-bound and contextual; business activities are subject to specific local and community conditions and terms. In the case where AI is deployed, this link with a territorial context may be less meaningful.

The High-Level Expert Group on Artificial Intelligence also recommends conducting a fundamental rights impact assessment.

In some cases, the use and deployment of AI can have high impact on social dynamics in certain territorial contexts, such as smart cities plans and predictive crime programmes. Similar to TRUSTEE, some predictive programs can be implemented through the deployment of AI whereby educational data in combination with energy data can be used to develop sustainability programs. There also cases in which AI functionalities would have a more limited impact as they are embedded in global products or services such as virtual assistants, autonomous cars, recruiting AI-based software.

As an example of TRUSTEE, AI can be implemented to do summations of volcanic activities on the moon, which is data obtained from the SPACE domain. These functions would not necessarily focus on a given socio-territorial community. While in the case with the educational data on TRUSTEE being used for predictive analytics of the consumption of energy in schools, can come much closer to the traditional HRIA cases, sustainability programs will indeed be able to impact whole communities (the education ecosystem) of how they utilise their energy, buildings and budgets, but they equally also can potentially impact a wide range of human rights – the privacy of the EDUCATION stakeholders (including vulnerable individuals such as students/children) but also the communities if there are any punitive actions as a result of identifying schools' energy waste.

In the second case where AI is deployed for purposes such as summarising or analysing data on the activities or state of the moon, the social impact is much limited and therefore would likely be focusing on specific individuals (perhaps the community from science and space domains) than society more generally. This difference has a direct effect on the structure and complexity of the framework and toolset that is proposed in this Deliverable. In both of these scenarios, the unique proposition in this Deliverable that will guide technical and non-technical stakeholders of TRUSTEE, is with regards to the socio-ethical and human rights assessment which will be ex ante in nature of the assessment carried and the greater focus on quantifiable risk thresholds. The ex-ante approach will be executed

at the TRUSTEE project design and development, instead of relying on ex post evaluation which then demands corrective policies that are typical of traditional HRIA³⁶⁹.

POTENTIAL RISKS OF ETHICAL AI

AI can be beneficial from both an economic and a social point of view. However, despite the wide adoption of AI, there are several negative impacts and ethical and legal concerns. AI should be implemented carefully to minimize or better avoid any unintentional negative consequences. The development and deployment of AI models is not a procedure that can be hurried³⁷⁰. It needs time to make sure that everything works as it should and there are no hidden problems. AI-based systems should be continually monitored and be frequently reviewed even after they are launched.

Ethical AI principles and frameworks should be put into practice and mechanisms for a strong governance and compliance should be created³⁷¹. Some of these ethical AI principles are trust, fairness, and privacy.

Firstly, AI models should be robust and consistently accurate. However, the quality of an AI model can be only as good as the training data that were used to create it. Therefore, a careful selection and preparation of data is of outmost importance. Many problems arise when data is not fully understood and not well documented.

One of the most common issues is when AI models are biased. This should not happen and depends highly on the data used to train them which can either reflect the bias that is already embedded in society or can be an inaccurate representation of the real world. AI models should be fair, safe and their decisions must not discriminate and cause no harm.

Furthermore, AI models should respect privacy and protect sensitive information. Stakeholders should understand what data is being collected and why as well as whether users have consented to its collection and use. Companies and organizations should not collect more data than what is required by the AI models. The risk of non-compliance with various privacy regulations is increased as the amount of data that is collected increases³⁷².

Another important problem is the transparency of data and the AI models. People need to understand the datasets and be able to trace the algorithms back to their training data. The AI models' behaviour should be fully understood to avoid any unexpected outcomes and help users trust them.

The decisions that an AI model makes are of critical importance especially in certain domains such as Healthcare. The absence of explanations related to how these decisions are made, and how the data are utilized, can lead to a lack of trust, which poses a significant risk for the deployment of AI models. The recommendation of an AI system can also be even life threatening if there is no explanation. In a similar manner, the same risk might also arise in TRUSTEE, as the users may not be able to trust the AI models developed in the platform if there is a lack of explanations and understanding about their operation. As it can be concluded, the presence of the XAI tool in TRUSTEE is essential for developing trust in the deployed AI models of the Platform. Explainability is an essential part of any transparent system and can reveal bias, provide fairness and permit accountability.

³⁶⁹ World Bank, Nordic Trust Fund (2013) Human Rights Impact Assessments: A Review of the Literature, Differences with other forms of Assessments and Relevance for Development. *World Bank and Nordic Trust Fund*, Washington, pp. 8–9.

³⁷⁰ R. Eitel-Porter, "Beyond the promise: implementing ethical AI," *AI Ethics*, vol. 1, no. 1, pp. 73–80, Feb. 2021, DOI: 10.1007/s43681-020-00011-6.

³⁷¹ Supra note 12.

³⁷² Supra note 12.

There is no doubt that Artificial Intelligence in the HEALTH domain “has enormous promises to improve healthcare”³⁷³. Nevertheless, the use of Artificial Intelligence systems needs to abide by requirements set out by the legislation but also meet ethical standards and practices. Even though there are no harmonised set of best practices or ethical requirements, several principles are widely referred to and expected to be implemented when Artificial Intelligence is used in the HEALTH domain³⁷⁴. As HEALTH is one of TRUSTEE use case, the ethical principles described below need to be taken into account and implemented when developing TRUSTEE but also after the end of the project. An ethical principle should be understood as a “statement of a duty or a responsibility in the context of the development, deployment and continuing assessment of AI technologies for health”³⁷⁵. Six ethical principles have been identified by the World Health Organisation (WHO) for the HEALTH domain.

1. *Autonomy*: autonomy refers to human decisions and consent of the data subject on both a data protection and ethical level. Decision-making is a critical issue in AI in the sense that decision-making can be or is transferred to machines, which goes against the principle of autonomy. When it comes to healthcare, human should remain in control of the systems and medical decisions. Additionally, human oversight must be assessed by the level of risks associated with an AI system. This ethical principle is inextricably linked to Article 22 GDPR, dealing with decision-making. In order to preserve autonomy at its fullest, a valid and informed consent must be given by the natural person on data protection level but also on an ethical level (e.g. consent to participate in research), the use of AI technologies for “experimentation or manipulation of humans in a healthcare system without a valid consent”³⁷⁶ should be forbidden. Autonomy connects with data protection, since the information to natural persons as well as some requirements for the purpose of processing health data (qualified as sensitive by Article 9 GDPR). Ethics are inextricably linked with data protection, especially in the HEALTH domain. As stated by the European Commission, “wherever you have the possibility to enhance the level of data protection afforded to your research subjects, you should apply such measures by default rather than just considering them or making them available as an optional extra.”³⁷⁷ On that note, “[o]ne of the best ways to mitigate the ethical concerns arising from the use of personal data is to anonymise them so that they no longer relate to identifiable persons”³⁷⁸. Therefore, in order to remove any risk inherent to ethical concern, all personal data related to HEALTH could be, if possible, anonymised locally before being brought to TRUSTEE.
2. *Human well-being, human safety and the public interest*: before any deployment of the AI technology, there must be a quality check to ensure safety, accuracy and efficacy on the one hand as well as measures allowing developers to provide for quality control and improvement.
3. *Transparency, explainability and intelligibility*: in order to build a trustworthy and ethical AI, the system should be intelligible and/or understandable to developers, users or regulators. Transparency is achievable

³⁷³ Naik N, Hameed BMZ, Shetty DK, Swain D, Shah M, Paul R, Aggarwal K, Ibrahim S, Patil V, Smriti K, Shetty S, Rai BP, Chlosta P and Somani BK (2022) Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? *Front. Surg.* 9:862322. doi: 10.3389/fsurg.2022.862322

³⁷⁴ Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.

³⁷⁵ Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.

³⁷⁶ Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.

³⁷⁷ European Commission, ‘Ethics and Data Protection’, 14 November 2018 <https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf>).

³⁷⁸ Ibid.

through information that means notably to explain and detail how the AI technology is designed, its limitations and the development of the algorithm. In regard to HEALTH, when the AI leads to decision-making, the information on how the decision was made should be published and the technology to be re-evaluated regularly.

4. *Responsibility and accountability*: responsibility is not yet clearly defined by the law as there is a proposal for a directive on AI liability. However, a “human warranty” is an assurance of responsibility and natural persons such as stakeholders are responsible for the efficiency and the conditions in which the AI system operates. Just as stated in Ethical Charter of the European Commission for Efficiency of Justice, as in the HEALTH domain there are sensitive information, the AI needs to rely on “certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment”³⁷⁹ – which should be achievable by TRUSTEE considering there are six use cases.
5. *Inclusiveness and equity*: healthcare should be accessible without bias or discrimination, also for public interest. Therefore, when building the platform and setting up the AI technology, one should consider different languages, ability and forms of communications worldwide, so it is as accessible as possible. Its access and the use must encompass any age, gender, socio-economic background, ethnicity, etc. The AI needs to prevent worsening or increasing any discrimination, whether it is between individuals or between groups of individuals³⁸⁰. Developers need to be particularly careful to unintended bias that may emerge from AI in the “design, implementation and use and the potential harm that biases can cause to individuals and society”³⁸¹. One way to mitigate the risk of unintended bias is to monitor and evaluate the AI system regularly before, during and after deployment.
6. *Sustainability*: WHO recommends that all actors pay particular attention to how the AI that was developed still answers to “communicated expectations and requirements in the context in which it is used”³⁸².

An impact assessment should be performed as it allows to better understand the potential consequences of the use of AI in the HEALTH domain. Performing an impact assessment is also beneficial in regard to technical information and improvement of decision-making and participates to fulfil the transparency requirement.³⁸³

CYBERSECURITY LEGAL FRAMEWORK AND REQUIREMENTS FOR INFORMATION SYSTEMS PROVIDERS

NIS DIRECTIVE AND EU CYBERSECURITY ACT

The NIS Directive aims to improve the cyber-security capabilities at a national level and foster better communication across EU Member States. Each Member State, transposing the NIS Directive, had to adopt a national strategy on achieving a high level of cyber-security for networks and information systems across essential services.

³⁷⁹ Ethical Charter of the European Commission for Efficiency of Justice : 5 principles relevant to use of AI for health, 2018

³⁸⁰ Ethical Charter of the European Commission for Efficiency of Justice : 5 principles relevant to use of AI for health, 2018

³⁸¹ Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.

³⁸² Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.

³⁸³ Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021.

In accordance with its Article 1(2)(d), it establishes security and notification requirements for *operators of essential services* and for *digital service providers*. Therefore, it is crucial to define whether TRUSTEE Partners, in the context of their role in the TRUSTEE project, should be perceived as one of these entities and if the requirements laid down in the NIS Directive apply to the TRUSTEE data space. The definitions of both types of concerned entities are indicated in Article 4 of the NIS Directive.

‘Operator of essential services’ (hereafter “OES”) means a public or private entity of a type referred to in Annex II, which meets the criteria laid down in Article 5(2)³⁸⁴. The types of entities, which may be relevant for TRUSTEE, listed in Annex II, are the entities from the digital infrastructure sector. These are: the providers of Internet Exchange Points, Domain-System-Service providers and Top-Level-Domain service providers. None of the TRUSTEE Partners represents this type of entity.

Moreover, the NIS Directive applies also to the entities defined in Article 4(6) of the Directive – “digital service providers”, defined as any legal person that provides a digital service. The digital service, as defined in Article 4(5) in conjunction with Article 1(1)(b) of the Directive (EU) 2015/1535 of the European Parliament and of the Council, means any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a receipt of services, which is of a type listed in Annex III to the NIS Directive: online marketplace, online search engine and cloud computing service.

Taking this definition into consideration, at present none of the TRUSTEE Partners represents a group of digital service providers in the meaning of the NIS Directive in the context of providing the TRUSTEE service, as at the moment none of the services in the context of TRUSTEE will be provided to the end users for remuneration. Moreover, none of the Partners will provide to the end user neither online marketplace, online search engine nor cloud computing services for remuneration.

Nevertheless, as a platform aiming to ensure the highest standards of cyber security and considering a cross border and cross sectoral usage of the platform, it is recommended that TRUSTEE fulfils the requirements posed by the NIS Directive and relevant national regulations and ENISA guidelines established for digital service providers.

The NIS Directive’s main objective is to introduce measures with a view to achieving a high common level of security of network and information systems within the Union to improve the functioning of the internal market (Article 1 NIS Directive). Being a directive, it had to be transposed by the Member States into their national legal systems. The Member States were requested to create a national strategy on the security of network and information system and provide obligations for operators of essential services and DSPs – digital service providers.

The NIS Directive applies to the legal entities providing cloud services, defined as ‘a digital service that enables access to a scalable and elastic pool of shareable computing resources’ (Article 4(19) NIS Directive).

According to Article 16 (1) NIS Directive, the cloud service providers are required to identify and take appropriate and proportionate technical and organizational measures to manage risks created by the security systems. In line with the mentioned Article, the level of security which should be ascertained has to be appropriate to the risks posed, and should take into account the following elements:

- a. the security of systems and facilities.
- b. incident handling.
- c. business continuity management.
- d. monitoring, auditing and testing; and
- d. compliance with international standards.

³⁸⁴ NIS Directive, Article 4(4).

These elements have been further interpreted in the European Commission Implementing Regulation.

Another obligation imposed on the cloud services provider in that regard is to prevent and minimize the impact of incidents affecting their network. According to the implementing regulation, ‘with regard to incident handling referred to in point (b) of Article 16(1) of Directive (EU) 2016/1148, the measures taken by the digital service provider shall include: (a) detection processes and procedures maintained and tested to ensure timely and adequate awareness of anomalous events; (b) processes and policies on reporting incidents and identifying weaknesses and vulnerabilities in their information systems; (c) a response in accordance with established procedures and reporting the results of the measure taken; (d) an assessment of the incident's severity, documenting knowledge from incident analysis and collection of relevant information which may serve as evidence and support a continuous improvement process.’ (Article 2(1) Implementing Regulation).

Moreover, it is required that the cloud service provider ensures appropriate business continuity management, i.e. “the capability of an organisation to maintain or as appropriate restore the delivery of services at acceptable predefined levels following a disruptive incident and shall include: (a) the establishment and the use of contingency plans based on a business impact analysis for ensuring the continuity of the services provided by digital service providers which shall be assessed and tested on a regular basis for example, through exercises; (b) disaster recovery capabilities which shall be assessed and tested on a regular basis for example, through exercises.” (Article 2(3) Implementing Regulation).

To fulfil the obligation of monitoring, auditing and testing, the Implementing Regulation indicates that there should be policies implemented which will define supervision, inspection and verification procedures. (Article 2(4) Implementing Regulation).

The compliance with all above-described points should be appropriately documented.

The NIS Directive establishes a framework according to which the Member States are required to:

- ensure that appropriate and proportionate Technical and organizational measures are defined and adopted by the DSPs,
- ensure that the DSPs act in a manner that prevents and minimizes the impact of incidents in case they occur,
- notify a competent authority or incident response team of any incident that has a substantial impact on the provision of cloud computing services without undue delay.

Additional, more relevant guidelines which allow for better understanding of the NIS Directive and provide recommendations and guidelines to the DSPs (so also to the cloud services providers) are introduced by ENISA.

Another European legal instrument having impact on the cloud computing services security is the Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification, which entered into force in 27 June 2019 (EU Cybersecurity Act). This act, apart from defining the role of ENISA, lays down the EU cybersecurity certification framework and the procedure for the creation of EU cybersecurity certification schemes, including the ICT products, services and processes. The scheme is currently under development under the leadership of ENISA. Once it is released, it will be further explored by the TRUSTEE project.

NIS 2 DIRECTIVE AND FUTURE DEVELOPMENTS

The NIS 2 Directive³⁸⁵ aims “to harmonise cybersecurity requirements and implementation of cybersecurity measures in different Member States. To achieve this, it sets out minimum rules for a regulatory framework and lays down mechanisms for effective cooperation among relevant authorities in each member state. It updates the list of sectors and activities subject to cybersecurity obligations and provides for remedies and sanctions to ensure enforcement”.³⁸⁶

NIS2 builds on the original NIS Directive, but it includes significant changes that aim to improve the resilience and incident response capacities of the public and private sectors and the EU.

Firstly, NIS2 has a broader scope of sectors and entities than those covered by the current NIS Directive, classifying them according to the groups of sectors:

- energy, transport, banking, financial market infrastructures, health, drinking water, waste water, digital infrastructure, ICT service management (B2B), public administration, space (all of them listed in Annex I and called “sectors of high criticality”),
- “other critical sectors”, listed in Annex II, such as, among others, digital providers³⁸⁷ and research³⁸⁸.

The entities recognised as relevant in the digital infrastructure sector are DNS service providers, TLD name registries, cloud computing service providers, data centre service providers, content delivery network providers, managed service providers, managed security service providers, providers of online marketplaces, of online search engines and of social networking services platforms, and trust service³⁸⁹ providers. They should be subject to a high degree of harmonisation at Union level due to their cross-border nature.³⁹⁰

The Directive recognises two main types of entities:

- *Essential entities*, among which there are qualified trust service providers, TLD domain name registries and DNS service providers (regardless of their size), the types of entities associated with the sectors of high criticality mentioned above, bigger than medium-size enterprise³⁹¹, public administration entities of central government as defined by a Member State in accordance with national law, any other entities from the sectors of high criticality and other critical sectors, which are identified by a member State as essential entities or operators of essential services identified as such according to NIS Directive.
- *Important entities*, i.e., entities listed in Annex I and II, which do not qualify as essential entities according to Article 3(1) NIS 2 Directive.

³⁸⁵ DIRECTIVE (EU) 2022/2555 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1772, and repealing Directive (EU) 2016/1148 (NIS 2 Directive) (OJEU L333, 27.12.2022, p. 80).

³⁸⁶ <https://www.consilium.europa.eu/en/press/press-releases/2022/11/28/eu-decides-to-strengthen-cybersecurity-and-resilience-across-the-union-council-adopts-new-legislation/>

³⁸⁷ Included entities: providers of online marketplaces, providers of online search engines and providers of social networking services platforms.

³⁸⁸ The entities considered here are research organisations.

³⁸⁹ NIS Directive, Article 6(24). ‘Trust service’ means a trust service as defined in Article 3, point (16), of Regulation (EU) No 910/2014 (eIDAS Directive), i.e. an electronic service normally provided for remuneration which consists of (a) the creation, verification, and validation of electronic signatures, electronic seals or electronic time stamps, electronic registered delivery services and certificates related to those services, or (b) the creation, verification and validation of certificates for website authentication; or (c) the preservation of electronic signatures, seals or certificates related to those services;

³⁹⁰ NIS 2 Directive, Recital 84.

³⁹¹ Under Article 2 of the Annex to Recommendation 2003/361/EC.

A network and information system shall be understood as digital data stored, processed, retrieved or transmitted for the purposes of their operation, use, protection and maintenance by elements such as any device or group of interconnected devices, one or more of which, pursuant to a programme, carry out automatic processing of digital data³⁹².

The Directive imposes direct obligations on "management bodies" concerning implementing and supervising their organization's compliance with the legislation.³⁹³ Failing to comply with those obligations might lead to potential fines and a temporary ban from discharging managerial functions.

NIS 2 Directive also includes a set of obligations requiring entities to implement cyber risk management measures, which include security risk mitigation requirements and third parties' due diligence³⁹⁴. The level of security of security of network and information systems should be appropriate to the risks posed, taking into account the state-of-the-art and relevant European and international standards. The measures shall include at least³⁹⁵:

- policies on risk analysis and information system security,
- incident handling³⁹⁶,
- business continuity³⁹⁷,
- supply chain security,
- security in acquisition, development and maintenance of network and information system,
- assessment of the effectiveness of implemented measures,
- cybersecurity training and basic cyber hygiene practices,
- procedures regarding the use of cryptography and, where appropriate, encryption, access control policies, asset management and HR security,
- the use of multi-factor authentication or continuous authentication solutions within the entity, where appropriate.

It must be taken into account that the providers of TRUSTEE system architecture may be recognised either as essential (as provider of trusted services) or as important (as research institutions) entities, therefore they should adhere, in all phases of the development of TRUSTEE and its further maintenance, to the NIS 2 Directive stemming obligations, which will be further transposed and established at the national level.

TRUSTEE should, however, be aware that the Commission will adopt implementing acts laying down the technical and the methodological requirements of the listed measures for providers of online marketplaces, trusted service providers, cloud computing service providers managed service providers and managed security service providers and may adopt such acts with regard to essential and important entities other than these just listed.

EIDAS REGULATION AND REQUIREMENTS

³⁹² NIS 2 Directive, Article 6(1)(c).

³⁹³ NIS 2 Directive, Article 20.

³⁹⁴ NIS2 Directive, Article 21.

³⁹⁵ NIS2 Directive, Article 21(2).

³⁹⁶ Incident report requirements impose new obligations on entities, establishing notification obligations in phases, including an initial notification within 24 hours of becoming aware of specific incidents or cyber threats (instead of simply "without undue delay" as in the NIS Directive), "intermediate" and "final" reporting obligations.

³⁹⁷ NIS 2 Directive, Article 21(2)(c). Such as backup management and disaster recovery, and crisis management.

In the context of eIDAS combined with SSI, it is relevant to further recognise impact of the GDPR and eIDAS Regulation on the design, development and implementation of such technical solution. The framework has been defined in the sections above – this section will address an impact of another regulation.

The Regulation on electronic identification and trust services for electronic transactions in the internal market (eIDAS Regulation) establishes the legal framework for electronic identification means and trust services, such as electronic signatures, seals, time stamps, documents, registered delivery services and certificate services for website authentication in the EU.³⁹⁸

The TRUSTEE system will incorporate a Self-Sovereign Identification management system, to allow for privacy over the identity. The aim is to provide a solution that allows the recipient entity to have cryptographically verifiable proof of the sender of the encrypted information. Furthermore, TRUSTEE solution will enable issuing, holding and verifying these SSI credentials and proofs that are eIDAS enabled. The eIDAS trust framework will support the development and implementation of the SSI concept over the eIDAS platform. TRUSTEE platform will showcase the deployment of blockchain technology that could leverage and benefit from the eIDAS regulatory framework. The application of eIDAS authentication will also enable onboarding of devices in a trusted and secure way while also guaranteeing the identification of the users of the TRUSTEE solution.

To be able to fulfil that, the TRUSTEE developers need to consider the legal framework set out in eIDAS Regulation for the eIDAS solutions.

eIDAS Regulation lays down the rules on electronic identification, i.e., mutual recognition, notification and cross-border use and further regulates trust services - electronic signatures, seals, timestamps and qualified services.

The eIDAS requires a mutual recognition of the eIDs in the EU. i.e., when an electronic identification using an electronic identification means³⁹⁹ and authentication is required to access a service provided by a public sector body online in one Member State, the electronic means issued in another Member State shall be recognized for the purposes of cross-border authentication for that that service online. However, under the conditions that the electronic means is included as an electronic identification scheme in a list published by the European Commission and the assurance level of this means corresponds to an assurance level equal or higher than the one required by the relevant public sector body to access the service in the country where the recognition is sought. The relevant public sector body must use the assurance level substantial or high in relation to accessing that service online.⁴⁰⁰

The levels of assurance established in the eIDAS regulation indicate the degree of trust and confidence in the system. It has to be always specified in the eIDs. Article 8 defines three assurance levels: low, substantial and high:

- a) *assurance level low shall refer to an electronic identification means in the context of an electronic identification scheme, which provides a limited degree of confidence in the claimed or asserted identity of a person, and is characterised with reference to technical specifications, standards and procedures related thereto, include technical controls, the purpose of which is to decrease the risk of misuse or alteration of the identity.*
- b) *assurance level substantial shall refer to an electronic identification means in the context of an electronic identification scheme, which provides a substantial degree of confidence in the claimed or asserted identity of a person, and is characterised with reference to technical specifications, standards and procedures*

³⁹⁸ Regulation (EU) No 910/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC.

³⁹⁹ eIDAS Regulation, Article 6.

⁴⁰⁰ eIDAS Regulation, Article 6.

related thereto, including technical controls, the purpose of which is to decrease substantially the risk of misuse or alteration of the identity.

- c) *assurance level high shall refer to an electronic identification means in the context of an electronic identification scheme, which provides a higher degree of confidence in the claimed or asserted identity of a person than electronic identification means with the assurance level substantial, and is characterised with reference to technical specifications, standards and procedures related thereto, including technical controls, the purpose of which is to prevent misuse or alteration of the identity.*⁴⁰¹

According to the Regulation, a *trust service* is an electronic service normally provided for remuneration which consists of:

- (a) the creation, verification, and validation of electronic signatures, electronic seals or electronic time stamps, electronic registered delivery services and certificates related to those services, or
- (b) the creation, verification and validation of certificates for website authentication; or
- (c) the preservation of electronic signatures, seals or certificates related to those services;⁴⁰²

The Regulation distinguishes trust services and qualified trust services. to be considered as qualified trust service, the service must meet a list of requirements set out in the Regulation.

The entities providing one or more trust services are defined as trust service providers⁴⁰³ (TSPs) The Regulation indicates obligations binding for those TSPs – among them the obligation to take appropriate technical and organizational measures to manage the risks posed to the trust services they provide.⁴⁰⁴ Measures shall be taken to prevent and minimise the impact of security incidents and inform stakeholders of the adverse effects of such incidents. The TSPs should also, inform the supervisory body or other relevant bodies (such as the national bodies responsible for information security or the data protection authorities) of any breach of security or loss of integrity that has a significant impact on the trust service provided or on the personal data maintained therein.

Where the breach of security or loss of integrity is likely to adversely affect a natural or legal person to whom the trusted service has been provided, the trust service provider shall also notify the natural or legal person of the breach of security or loss of integrity without undue delay.⁴⁰⁵

The eIDAS Regulation identifies three types of electronic signatures, which are going to be considered when designing the architecture of TRUSTEE data space with respect to the access management:

- (Basic) e-Signature⁴⁰⁶

⁴⁰¹ eIDAS Regulation, Article 8(2).

⁴⁰² eIDAS, Article 3(16).

⁴⁰³ eIDAS Regulation, Article 3(19) and (29).

⁴⁰⁴ eIDAS Regulation, Article 19(1).

⁴⁰⁵ eIDAS Regulation, Article 19(2).

⁴⁰⁶ eIDAS Regulation, Article 3(10). This type of the signature refers to any kind of signature, made in an online environment by a natural person, that manifests the intention of the signatory to be bound by the contents of the signed document. In practice, this could be clicking a button or checking a box.

- Advanced e-Signature^{407 408}
- Qualified e-Signature⁴⁰⁹. The Regulation sets out strict requirements for the qualified electronic signature creation devices. These devices shall adopt technical and procedural means to ensure that:
 - a) The confidentiality of the electronic signature creation data used for electronic signatures is reasonably assured.
 - b) The electronic signature creation data used for electronic signature creation can practically occur only once.
 - c) The electronic signature creation data used for electronic signature creation cannot, with reasonable assurance, be derived and the electronic signature is reliably protected against forgery.
 - d) The electronic signature creation data used for electronic signature creation can be reliably protected by the legitimate signatory against use by others.⁴¹⁰

The use of qualified electronic signatures and corresponding certificates invokes an additional layer of assurance and trust that results in a special legal effect that is recognised by the courts in the EU.⁴¹¹

Moreover, eIDAS Regulation has an impact also on the technologies making use of blockchain. Having considered the definition of electronic documents, it has to be noted that the data in blockchains could qualify as the electronic documents, as this data represents “any content stored in electronic form, in particular text or sound, visual or audiovisual recording”⁴¹². Hence, this data may have a legal effect, even though provided only in electronic form.

⁴⁰⁷ eIDAS Regulation, Article 3(11). This refers to an electronic a signature that meets the requirements set out in Article 26 of the eIDAS Regulation namely, it must be uniquely linked to the signatory; capable of identifying the signatory; created using electronic signature creation data that the signatory can use under his sole control and with a high level of confidence; and linked to the data signed therewith in such a way that any subsequent change in the data is detectable.

The Commission laid down further specifications for formats of advanced e-Signatures to be recognised by public sector bodies in the Decision (EU) 2015/1506. If these specifications are satisfied by the e-Signature, it may be assumed that this signature corresponds with the requirements of an Advanced e-Signature.

⁴⁰⁸ eIDAS Regulation; Article 27 (5), and Commission Implementing Decision (EU) 2015/1506 of 8 September 2015 laying down specifications relating to formats of advanced electronic signatures and advanced seals to be recognised by public sector bodies pursuant to Articles 27(5) and 37(5) of Regulation (EU) No 910/2014 of the European Parliament and of the Council on electronic identification and trust services for electronic transactions in the internal market.

⁴⁰⁹ eIDAS Regulation, Article 3(12). This type of eSignature is an “advanced electronic signature that is created by a qualified electronic signature creation device, and which is based on a qualified certificate for electronic signatures”. Therefore, a qualified electronic signature must meet the requirements in Article 26 (identified above), originate from a qualified electronic signature creation device and be based on a qualified certificate for electronic signatures. Annex I of the eIDAS Regulation includes the requirements for the qualified certificate for electronic signatures. The electronic signature creation device is defined as a “configured software or hardware used to create an electronic signature” (see eIDAS, Article 3(22) and (23)).

⁴¹⁰ eIDAS Regulation, Annex II.

⁴¹¹ Article 25 of the eIDAS Regulation states that a qualified eSignature has the equivalent legal effect of a handwritten signature. Further, it is admissible as an evidence in legal proceedings. This Article holds that eSignatures cannot be denied legal effect solely because “it is in an electronic form or does not meet the requirements for a qualified electronic signature”.⁴¹¹ Additionally, a qualified e-Signature ensures appropriate level of data integrity, data security and the secure identification of the signatory.

⁴¹² eIDAS Regulation, Article 3(45).

Based on the above, the literature provides a list of legal and regulatory requirement which should be considered when implementing eIDAS over SSI, and examples of these⁴¹³ have been considered in the legal requirements list presented in Part III of the Deliverable.

DIGITAL SERVICES ACT AND REQUIREMENTS FOR INTERMEDIARY SERVICES

The Digital Services Act (hereinafter “DSA”) aims to harmonise the rules regulating the liabilities and obligations of online intermediaries and platforms to create a safe digital space and a high level of protection for its participants while also creating legal certainty to providers of these services on their responsibilities. While affirming the objectives of the e-Commerce Directive, the DSA seeks to impose new measures that are centered on consumer protection in the digital environment.

The three main pillars of this Regulation are:

- Framework set for the liability exemptions for intermediary service providers, which is almost transferred as it is from the E-Commerce Directive⁴¹⁴;
- Rules governing asymmetric due diligence obligations for intermediary service providers, according to the role, size and impact of the provided services.
- Rules governing the implementation and enforcement of the Act.⁴¹⁵

The place of establishment of the intermediary service providers is of no relevance to the application of the Act, as long as the recipients of the service are located in the Union.⁴¹⁶

The definition given for the intermediary services in the Act are two-tiered. First, the DSA categories them within the general definition of “information society services” as accepted under the e-Commerce Directive, which are services that are:

- “provided at a distance” (*i.e. the service is provided without the parties being simultaneously present*),
- “by electronic means” (*i.e. the service is sent initially and received at its destination by means of electronic equipment for the processing and storage of data*) and,
- “at the individual request of a recipient of services” services (*i.e. the service is provided through the transmission of data on individual request*).⁴¹⁷
- “normally provided for remuneration”⁴¹⁸ - such services are not limited to the ones giving rise to online contracting but also, in so far as they represent an economic activity, cover services which are not remunerated by those who receive them (*i.e. services offering online information or commercial communications; services providing tools allowing for search, access and retrieval of data; services in hosting information provided by a recipient of the service*).⁴¹⁹ Instead, the services may be

⁴¹³ Michael Kuperberg, Blockchain-Based Identity Management: A Survey From the Enterprise and Ecosystem Perspective, IEEE Transactions on Engineering Management, vol. 67, No. 4, November 2020, p. 1015.

⁴¹⁴ Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'). The DSA will not affect the application of this Directive.

⁴¹⁵ DSA, Article 1.

⁴¹⁶ DSA, Article 2(1).

⁴¹⁷ Directive 2015/1535 of the European Parliament and of the Council of 9 September 2015 laying down procedure for the provision of the information in the field of technical regulations and of rules on Information Society services, Article 1(1)(b).

⁴¹⁸ Ibid.

⁴¹⁹ E-Commerce Directive, Recital 18.

remunerated, for instance, by the incomes generated by advertisements posted on a website⁴²⁰, the costs of the performance of a service for the purposes of advertising the goods sold and services provided which are incorporated in the price of goods and services.⁴²¹

The DSA later divides intermediary services into three different categories:

- mere conduit service: “transmission in a communication network of information provided by the recipient of the service or the provision of access to a communication network”⁴²²;
- caching service: “transmission in a communication network of information provided by a recipient of the service, involving the automatic, intermediate and temporary storage of that information, performed for the sole purpose of making more efficient the information’s onward transmission”⁴²³;
- hosting service: services for the “storage of information provided by, and at the request of, a recipient of the service”⁴²⁴, such as cloud and webhosting services, also including online platform services (etc. online marketplaces, social media platforms, app stores).

TRUSTEE is intended to become an online solution to facilitate the trustworthy access and re-use of data, and to provide services to search and discover data. With this in mind, what needs to be clarified first is whether the solution aimed to be developed fits into the definition of an “information society service” as given by the E-Commerce Directive and the DSA.

For this moment, considering that the TRUSTEE platform and the services to be provided therein are developed within the context of a research project, funded by public resources, and are to be freely accessible by the partners, there does not seem to be an engagement in an economic activity as required by the E-Commerce Directive in its definition of an established service provider.⁴²⁵ Any vision as to the commercialization of and profiting from the TRUSTEE platform through making it available to users are also not currently envisaged, as well as financing the service via any advertising or sponsorship revenues.

However, upon any change in circumstances, which would require the consideration of commercialization, engagement in an economic activity and any type of remuneration for providing the services that would lead to defining the TRUSTEE platform service as an information society service, TRUSTEE partners might be required to abide by several obligations stipulated under the DSA. Further clarification of the following should follow:

- How the TRUSTEE platform would fit under the “hosting service”, more specifically the “online platform” definition within the intermediary services definition governed by the DSA.⁴²⁶
- Whether and how, for the provision of the platform services, a contract would need to be signed between TRUSTEE service providers and TRUSTEE users as the recipients of the service. Not only that Article 9 of e-Commerce Directive requires member states to ensure that their legal system allows

⁴²⁰ CJEU, C-291/13 Papasavvas EU:C:2014:2209, para. 30

⁴²¹ Lubasz D., Namyslowska M., ‘EU Digital Law, Article 5 General information to be provided’, 2020, p. 418, referring to CJEU, C-484/14 McFadden EU:C:2016:689, para. 42 and case-law cited within.

⁴²² DSA, Article 3(g)(i).

⁴²³ DSA, Article 3(g)(ii).

⁴²⁴ DSA, Article 3(g)(iii).

⁴²⁵ According to Article 2(a)(b) of the e-Commerce Directive, a service provider is “any natural or legal person providing an information society service”, and an established service provider is one who “effectively pursues an economic activity using a fixed establishment for an indefinite period” in the EU.

⁴²⁶ The DSA, under Article 3(i), defines online platforms as hosting services that “at the request of a recipient of the service, stores and disseminates information to the public, unless that activity is a minor and purely ancillary feature of another service or a minor functionality of the principal service and, for objective and technical reasons, cannot be used without that other service, and the integration of the feature or functionality into the other service is not a means to circumvent the applicability of this Regulation.”

for contracts to be concluded by electronic means, the DSA also brings forth obligations with respect to the terms and conditions governing the contractual relationship between the provider of intermediary services and the recipients of the service. The terms and conditions to be drafted for the platform would have to be in clear, plain, intelligible, user-friendly and unambiguous language, and shall be publicly available in an easily accessible and machine readable format.

- Whether the TRUSTEE system developers/providers would plan to include on the platform any commercial communication such as advertisements or sponsorship. If this would be the case, recipients of the services of the TRUSTEE platform would need to be able to, for each specific advertisement presented, identify that the information is an advertisement in a clear, concise and unambiguous manner and in real time, on whose behalf the advertisement is presented and who paid for that advertisement.⁴²⁷

Whether the TRUSTEE solution would also allow consumers to conclude distance contracts with traders through the platform, in the form of a marketplace for the trading of data. If this is to be decided upon at the Consortium level, an evaluation as to how and whether additional obligations for such activity, as envisaged under Section 4 of the DSA, would be applicable to the TRUSTEE platform.

P2B REGULATION

The P2B Regulation⁴²⁸ introduces new rules that are directly applicable in each Member State to online intermediation services and online search engines that aim to connect EU businesses and professional websites with EU consumers. It applies to online intermediation services and online search engines provided, or offered to be provided, to business users and corporate website users, respectively, that have their place of establishment or residence in the EU and that, through those online intermediation service or online search engines, offer goods or services to consumers located in the EU, irrespective of the place of establishment or residence of the providers of those services and irrespective of the law otherwise applicable.⁴²⁹

The P2B Regulation defines “online intermediation services” as services which meet all of the following requirements⁴³⁰:

- a) they constitute information society services⁴³¹

⁴²⁷ DSA, Article 26(1).

⁴²⁸ Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services, OJ L 186/57.

⁴²⁹ P2B Regulation, Article 1(2).

⁴³⁰ P2B Regulation, Article 2(2).

⁴³¹ Within the meaning of point (b) of Article 1(1) of Directive (EU) 2015/1535 laying down a procedure for the provision of information in the field of technical regulations and of rules on Information Society services. According to this provision, the Information Service is a service provided normally for remuneration, at a distance, by electronic means and at the individual request of a recipient of services. For the purposes of this definition: (i) ‘at a distance’ means that the service is provided without the parties being simultaneously present; (ii) ‘by electronic means’ means that the service is sent initially and received at its destination by means of electronic equipment for the processing (including digital compression) and storage of data, and entirely transmitted, conveyed and received by wire, by radio, by optical means or by other electromagnetic means; (iii) ‘at the individual request of a recipient of services’ means that the service is provided through the transmission of data on individual request.

It should be further stressed that remuneration does not necessarily mean payment of money, but may be also understood as provision of personal data or other data. This includes the situations where: (a) the provider of a service requests and the end-user knowingly provides personal data or other data directly or indirectly to the provider; (b) the end-user allows access to information without actively supplying it, such as personal data, including the IP address, or other automatically generated information, such as information collected and

- b) they allow business users and corporate website users to offer goods or services to consumers located in the EU, with a view to facilitating the initiating of direct transactions between those business users / corporate website users and consumers, irrespective of where those transactions are ultimately concluded
- c) they are provided to business users on the basis of contractual relationships between the provider of those services and business users which offer goods or services to consumers.

The “provider of online intermediation services” is defined as any natural or legal person which provides, or which offers to provide, online intermediation services to business users.⁴³²

The P2B Regulation imposes several obligations on the providers of online intermediation services, listed below.

Transparency obligations are the main focus point of P2B Regulation. The Regulation requires that service providers are clear and transparent in their terms and conditions (‘T&Cs’).

The T&Cs have to:

- be drafted in plain and intelligible language⁴³³
- be easily available to the business users in all phases of their commercial relationship with the service provider, including the pre-contractual phase⁴³⁴
- include obligatory information such as, in particular:
 - description of the grounds on which the provider of online intermediation service may decide to suspend, terminate or otherwise restrict the use of its services⁴³⁵ and the statement of reasons for a decision to restrict or suspend the provision of the services to a given business user in relation to the goods and services offered by that business user or a decision to terminate the provision of the whole of the services – the business user shall have the right to clarify the facts and circumstances in the framework of internal complaint-handling process and the platform will also have to reinstate sellers if a suspension was made in error;⁴³⁶
 - all relevant information relating to the access to and functioning of their internal complaint-handling system⁴³⁷ and mediation⁴³⁸
 - general information concerning the overall effects, if any, of the terms and conditions on the ownership and control of intellectual property rights of business users⁴³⁹
- disclose a range of business practices, by including:
 - a description of any differential treatment and ranking of goods and services offered by online platforms and search engines themselves (or professional users controlled by them), in comparison to those offered by third-party business users⁴⁴⁰;

transmitted by a cookie; (c) the service provider is paid by a third party and not by the service recipient - The concept of remuneration should therefore also encompass situations in which the end-user is exposed to advertisements as a condition for gaining access to the service, or situations in which the service provider monetises personal data it has collected. (Recital 16 of the Directive 2015/1535).

⁴³² P2B Regulation, Article 2(3).

⁴³³ P2B Regulation, Article 3(1)(a).

⁴³⁴ P2B Regulation, Article 3(1)(b).

⁴³⁵ P2B Regulation, Article 3(1)(c).

⁴³⁶ P2B Regulation, Article 4(1)(2)(3).

⁴³⁷ P2B Regulation, Article 11(3).

⁴³⁸ P2B Regulation, Article 12(1).

⁴³⁹ P2B Regulation, Article 3(1)(e).

⁴⁴⁰ P2B Regulation, Article 7(1).

- a description of any ancillary goods or services that the online platform may itself offer as a complement to the goods or services of its professional users is also required⁴⁴¹;
 - information about the legal, economic or commercial consideration for any restriction of the ability of professional users to offer their goods or services under different terms through other channels (*Best Price Clause*)⁴⁴² - this applies, for instance, where the provider precludes business users from offering goods and services on better conditions or at lower prices on their own website or through other online intermediation services;
 - information under which conditions a business user can terminate its contractual relationship with the online intermediation service⁴⁴³;
 - a description of the access or absence thereof to any information provided or generated by business user where the online intermediation service maintains such access after the end of the contractual relationship⁴⁴⁴;
 - a description of the technical and contractual access (or its absence) of business users to any personal data or other data (or both) which business users or consumers provide for the use of the services or which are generated through the provision of the services - in other words, service providers must disclose what data they collect, and how they use it – and in particular how such data is shared with other business partners they have. Where personal data is concerned, the rules of the GDPR apply.⁴⁴⁵
- Make the identity of the its business users clearly visible (and may not prevent this)⁴⁴⁶

If the T&Cs do not comply with these requirements, they may be considered null and void (and non-enforceable).⁴⁴⁷

With regards to TRUSTEE services, it is not excluded that TRUSTEE could be considered as online intermediation services, enabling business users (data providers) to offer their goods and services (data in their possession) to the consumers (Consumers) for a remuneration (e.g. in the form of the possibility to access the data they provide in exchange).

B2B AND B2C DATA SHARING PRINCIPLES AND CONTRACTUAL ARRANGEMENTS

B2B DATA SHARING PRINCIPLES AND MODELS

The series of communications released by the European Commission on 25th April 2018 relates to data through AI and the Internet of Things (IoT). The Communication establishes five main guiding principles to consider when drafting a data sharing agreement:

- (1) Transparency: “the relevant contractual agreements should identify in a transparent and understandable manner (i) the persons or entities that will have access to the data that the product or service generates, the type of such data, and at which level of detail; and (ii) the purposes for using such data.”
- (2) Shared value creation: “the relevant contractual agreements should recognize that where data is generated as a by-product of using a product or service several parties have contributed to creating the data.”

⁴⁴¹ P2B Regulation, Article 6.

⁴⁴² P2B Regulation, Article 10.

⁴⁴³ P2B Regulation, Article 8(b).

⁴⁴⁴ P2B Regulation, Article 8(c).

⁴⁴⁵ P2B Regulation, Article 9.

⁴⁴⁶ P2B regulation, Article 3 para 5; Q&A P2B Regulation, MEMO/19/1169, https://ec.europa.eu/commission/presscorner/detail/en/MEMO_19_1169.

⁴⁴⁷ P2B Regulation, Article 3(3).

- (3) Respect for each other's commercial interests: "the relevant contractual agreements should address the need to protect both the commercial interests and secrets of data holders and data users."
- (4) Ensure undistorted competition: "the relevant contractual agreements should address the need to ensure undistorted competition when exchanging commercially sensitive data."

Therefore, the contractual agreements must comply with EU competition law, namely Article 101 and 102 TFEU.

- (5) Minimised data lock-in: "Companies offering a product or service that generates data as a by-product should allow and enable data portability as much as possible. They should also consider, where possible and in line with the characteristics of the market they operate on, offering the same product or service without or with only limited data transfers alongside products or services that include such data transfers."

Three models proposed by the Commission⁴⁴⁸:

- An Open Data approach: under this approach, the idea would be to have very little restrictions and barely any remuneration.
- Data monetisation on a data marketplace: the working group indicates that this approach of data monetisation or trading would not be suitable in all contexts. Indeed, it points out that for the mechanism to be applicable it would imply "(1) limited risks of illicit use of the data in question, (2) the data supplier has grounds to trust the (re-)user, or (3) the data supplier has technical mechanisms to prevent or identify illicit use."⁴⁴⁹
- Data exchange in a closed platform: in this model, the data can be provided against financial monetisation or added-value services. Under this approach, the Working group considers it would give a more comprehensive solution for more stable data partnerships and a better control over data usage. However, data competition law would apply to exclusive data sharing agreement.

TERMS OF DATA SHARING AGREEMENTS

When it comes to data sharing, contract law and the attached principle of freedom of contract apply. Unlike B2C relationships that are already subject to several regulations, B2B relationships rely mainly on freedom of contract and regulations that apply, without being specific to that type of agreement⁴⁵⁰. "[M]odel contract terms for different types of data sharing agreements and for some sectors or types of data sharing are already being developed"⁴⁵¹ without necessarily specifying what measures are specifically being developed. The documents provided by the Support Centre for data sharing may be used as model for contractual agreements, the Support Centre is funded by the European Commission to enhance and facilitate data sharing⁴⁵². In their Report on collected model contract terms, the Support Centre aims to "to provide a range of model contracts, which can be searched and used as templates by relevant stakeholders to license data, along with the right tools to easily find, select and customise a

⁴⁴⁸ Commission Staff Working Document, Guidance on sharing private sector data in the European data economy *accompanying* Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions "Towards a common European data space", SWD(2018) 125 final, p. 5

⁴⁴⁹ Ibid.

⁴⁵⁰ Begona Gonzales Otero, Evaluating the EC Private Data Sharing Principles: Setting a Mantra for Artificial Intelligence Nirvana ?, 10 (2019) JIPITEC 66 para 1. p.76

⁴⁵¹ Commission Staff Working Document, Guidance on sharing private sector data in the European data economy *accompanying* Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions "Towards a common European data space", SWD(2018) 125 final, p. 6

⁴⁵² [s://copa-cogeca.eu/img/user/files/EU%20CODE/EU_Code_2018_web_version.pdf](https://copa-cogeca.eu/img/user/files/EU%20CODE/EU_Code_2018_web_version.pdf) (last visited 6 October 2020). 273 Support Centre for Data Sharing website, <https://eudatasharing.eu/homepage>.

suitable licence”⁴⁵³. This document is aiming at using a “standardised classification system”⁴⁵⁴ by analysing different characteristics:

- (i) Provisions relating to the content or nature of the data
- (ii) Commercial/business related provisions
- (iii) Provisions relating to control, ownership and usage rights
- (iv) General provisions
- (v) Provisions related to the service providing the data (if any).

This document could be used as reference base to draft tailor-made data sharing and/or data licensing agreement for TRUSTEE since this report “can be considered as a soft and non-regulatory form of intervention”⁴⁵⁵.

In addition, the Staff Working Document provides insightful considerations⁴⁵⁶ – again non-binding – in the drafting or negotiation of data usage agreements (the following list is not exhaustive):

- Data made available: there is no specific definition of what “data” is, therefore it is advised to contractually define what “data” refers to whether it is a data asset purchase, a data stream or a data as a service⁴⁵⁷. The parties to the contract should describe the data they share “as concretely and precisely as possible”, the quality level planned for the data, compliance with legal obligations regarding access or transfer of the data and compliance with the relevant data protection legislation⁴⁵⁸.
- Access and (re-)usage of the data: the contractual provisions should define the rights to access, to (re)use, to distribution, sub-licensing and the conditions attached in a clear, transparent and understandable manner. The provisions should also include a detailed description of how the data will be used, to what extent it may be used and what usage is prohibited and confidentiality and non-disclosure rules. The usage should also include intellectual property rights, as well as authorizations for commercial or non-commercial purposes.
- Technical measures for data access and sharing and security measures: technical and security aspects for data access or exchange, IT security requirements, service levels for support, ensuring the protection of the data from threats, including – but not limited to – theft, misuse, technical problem and human error. In case security measure are not correctly provided for or implemented, it may lead to liability issues.
- Liability: refers to the responsibility of one party for harm or damage cause to another party, which may be a cause for compensation, financially or otherwise, by the former to the latter⁴⁵⁹. Every actor in the

⁴⁵³ Support Centre for Data Sharing, B.1 – Report on collected model contract terms Support Centre for data sharing. The report is available at: https://eudatasharing.eu/sites/default/files/2019-10/EN_Report%20on%20Model%20Contract%20Terms.pdf (last visited 5 October 2020).

⁴⁵⁴ Ibid. 6.

⁴⁵⁵ Ibid.

⁴⁵⁶ Commission Staff Working Document, Guidance on sharing private sector data in the European data economy *accompanying* Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions “Towards a common European data space”, SWD(2018) 125 final, p. 6

⁴⁵⁷ The roll-out of the EC data strategy will set up the foundation for enabling modern data management and the provisioning of services, including Data as a Service (DaaS), through the data platform. The European Commission Cloud Strategy is available at: https://ec.europa.eu/info/sites/info/files/ec_cloud_strategy.pdf (last visited 11 October 2020).

⁴⁵⁸ Commission Staff Working Document, Guidance on sharing private sector data in the European data economy *accompanying* Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions “Towards a common European data space”, SWD(2018) 125 final, p. 6

⁴⁵⁹ European Commission, ‘Commission Staff Working Document Liability for Emerging Digital Technologies Accompanying the Document Communication from the Commission to the European Parliament, the European

chain of the data sharing agreement should have their liability covered and expressly defined, these rules are left to the discretion of the parties, except for national mandatory liability rules.

- Parties' rights: the contract should provide for parties' rights including intellectual and industrial property rights, trade secrets, sensitive commercial information. The contract should include the right to perform audits.

Nevertheless, being a soft law solution, above principles have been considered in the initial list of the legal requirements for TRUSTEE. In case of any changes which would require a change of approach, the requirements will be modified accordingly and will be reflected in the Deliverable 2.2 and 2.3.

BLOCKCHAIN SERVICES – RELEVANT CONSIDERATIONS FOR BLOCKCHAIN PROVIDERS AND USERS

BLOCKCHAIN

Blockchain technology is an inevitable part of TRUSTEE architecture. TRUSTEE aims to “showcase the deployment of blockchain technology that could leverage and benefit from the eIDAS regulatory framework”, propose “enhancements in Self-sovereign technologies with the use of Blockchain for the decentralised identifiers”⁴⁶⁰, as well as to improve accountability for actions made on and via TRUSTEE by implementation of an Accountability Transaction Recorder (ATR). The ATR “will utilise different technological components for recording every single transaction of the users with the data while permitting private transactions where the content of the transaction can only be visible to authorized users as in the distributed ledger there are no administrative privileges.”⁴⁶¹ Furthermore, Security and Trust Manager, developed in T4.6, aims also to apply smart contracts.

This section tackles the interplay of the blockchain technology with law.

It needs to be understood that blockchain is a technology used for processing the data, but not data processing operation itself. Blockchain is a digital decentralised ledger technology. It is “distributed, shared and encrypted database that serves as an irreversible and incorruptible public repository of information”.⁴⁶² It is also developed chronologically, by adding new blocks, and may be defined by the following aspects:

- Sharing and decentralisation – copies of blockchain exist in parallel on different locations/machines
- Transparency – recorded data can be seen by all participants
- Lack of intermediation – there is no central entity deciding, all decisions are made in a form of a consensus between participants
- Irreversibility / tamper-proof – data once recorded cannot be removed or altered

The blockchain can be public or private.

- Public blockchain may be accessed by all, worldwide. In other words, transactions can be recorded, blocks validated and blocks' copies accessed by anyone.

Council, the Council, the European Economic and Social Committee and the Committee of the Regions Artificial Intelligence for Europe' SWD/2018/137 final, 2 footnote 1

⁴⁶⁰ TRUSTEE Grant Agreement, Part B, p. 3.

⁴⁶¹ TRUSTEE, Grant Agreement, Part B, Description Of The Action (Part A), p. 9.

⁴⁶² Aaron Wright and Primavera De Filippi, 'Decentralized Blockchain Technology and the Rise of Lex Cryptographia' (Social Science Research Network 2015) SSRN Scholarly Paper ID 2580664 2.

- Private blockchain can be a closed system, with access rights restricted to a specific group of entities (consortium) or a unique actor, who oversees participation and validation. If a new participant should join the network, this action has to be approved by the consortium, as the access to the network is enabled to a predefined group of actors. Moreover, different sets of rights (to read, to write or to administer) can be assigned individually. One of the examples of this type of blockchain is the one used in TRUSTEE – Hyperledger, being an example of “Blockchain 3.0”, related to the blockchains outside of the finance domain. These types of blockchain solutions are focused on decentralised, autonomous units that interact based on the smart contracts common for them. , and a Hyperledger Fabric is one of the main examples of it. Private blockchain is not decentralised and entities that operate the blockchain are enabled to make decisions concerning the network. This type of network allows for quality reviews and control of the data on the ledger, as well as allows for suspension, rectification or termination of the blockchain, which makes it mutable / reversible.

In the context of blockchain, one can distinguish three main actors:

- Participants – having the right to write on the chain and to decide to share the data for validation by the miners;
- Miners, validating transactions submitted by participants and not involved in the object of these transactions;
- Accessors – entities having “the right to read and hold a copy of the chain”⁴⁶³.

It is important for the users to understand what type of network they are supposed to use.

LEGAL ISSUES CONCERNING THE BLOCKCHAIN TECHNOLOGY

The EU does not offer a unified, pan-European blockchain regulation. Therefore, there are several legal frameworks that apply to this technology and several questions stemming from them that have to be monitored and addressed by the entities applying it. The below section includes some of them, having relevance for the TRUSTEE platform.

BLOCKCHAIN AND THE GDPR INTERPLAY

The GDPR applies when blockchain includes personal data (also encrypted/pseudonymised personal data). The GDPR however does not regulate this particular technology itself, but it sets out rules on how different entities use this technology when processing personal data.

With regards to personal data, a blockchain may involve two categories:

- Identifiers of participants’ and miners’ – they are in possession of a public key, which allows to identify of the issuer and receiver of a transaction
- Additional data that are included in the transaction - it may be information from a document or a specific database, concerning an individual other than participant or miner

If it includes such data, then GDPR compliance is required.

Technically speaking, “private” blockchains, due to the lack of their decentralisation and shared validation, should not pose any significant issues in terms of their compliance with the GDPR, as they are just distributed databases and their management allows for more GDPR-compliant approach.

⁴⁶³ [Blockchain and the GDPR: Solutions for a responsible use of the blockchain in the context of personal data | CNIL](#)

Points that require a particular attention from the data protection perspective are:

- Subcontracting obligations in the context of collaboration with subprocessors
- Adherence to the rules governing international transfers of personal data – transfers of data outside of the EU may become problematic, as participants and miners can be located in countries outside of the EU. To find the solution for that, permissioned or private blockchain could be considered, which could include appropriate safeguards, i.e. standard contractual clauses, codes of conduct, certification or binding corporate rules
- Adherence to privacy by design and by default principle – it should be always considered if using blockchain is really necessary;
- Applicability of the data protection impact assessment
- Storage limitation, data minimization and purpose limitation principles – reply to those principles very much depends on the type of personal data:
 - o in the case of identifiers of participants and miners, it seems that due to the architecture of blockchain these data are always visible and cannot be minimized and the storage period is adhered to the blockchain's existence.
 - o Additional data stored on blockchain, which may belong to data subjects other than participants and miners, it is recommended to process personal data outside of the blockchain and to store on blockchain only a proof of existence of the data – hash from a keyed hash function on the data, a ciphertext of the data or commitment of the data, or at least encrypt the data, to ensure a high level of confidentiality.
- Exercise of data subject's rights:
 - o The right to information, access and portability do not seem to be problematic, as blockchain allows for transparency and includes information about the history of the transaction.
 - o With regards to right to erasure, right to rectification and right to object to processing the situation is different. In this context, considering that the data are irreversibly written in the chain, it is important to avoid storing personal data in cleartext. "When the data recorded on the blockchain is a commitment, a hash generated by a keyed- hash function or a ciphertext obtained through "state of the art" algorithms and keys, the data controller can make the data practically inaccessible, and therefore move closer to the effects of data erasure."⁴⁶⁴
 - o Before writing a programme, e.g. before implementing a smart contract, data subjects' rights should be taken into account, and restriction of processing and human intervention should be enabled.⁴⁶⁵
- Security requirements:
 - o it is recommended to set a number of technical and organizational measures to minimize the impact of the failure of the algorithm, which may appear in a form of publication of a vulnerability on a cryptographic solution with regards to the security of transaction, and, e.g. to allow for replacing the algorithm after vulnerability identification.
 - o Confidentiality and security of secret keys is crucial and should be ensured by the data controllers (participants).
- Roles according to the GDPR – due to the technical construction of a blockchain, it is difficult to define who is data controller and who is a data processor. According to the guidance issued by CNIL, In the context of TRUSTEE, it seems that TRUSTEE system providers should be considered as miners, whereas Data Providers and Consumers, who define the purposes and the means (data format, use of blockchain technology) should be considered as participants, and hence – data controllers.⁴⁶⁶
In the personal data protection regime, smart contract developers, processing personal data for the data controller (in TRUSTEE - Data Provider and Consumer), are considered processors of the data.⁴⁶⁷ Additionally, also miners can be considered data processor, since they follow data controllers' instructions in the context of

⁴⁶⁴ [*Blockchain \(cnil.fr\)](#), p. 8

⁴⁶⁵ Ibid, p. 10

⁴⁶⁶ [*Blockchain \(cnil.fr\)](#), p. 1

⁴⁶⁷ [*Blockchain \(cnil.fr\)](#), p. 3.

checking compliance of transaction with technical requirements. This relationship would require a data processing agreement.⁴⁶⁸

LEGAL COMPLIANCE AND LIABILITY

Due to the fact that there is no entity that fully owns the blockchain due to its decentralised nature, it is difficult to determine the responsibility for legal compliance within the chain.⁴⁶⁹ The degree of difficulty of determining it depends on whether the blockchain is permissioned or permissionless. In permissionless blockchains, several actors located in various jurisdictions govern it. *Where a global network is controlled by a centralised legal entity (such as an online intermediary platform) this is not necessarily a problem. Identifying a regulatory access point is, however, more complicated where there is no centralised legal entity responsible for the network.*⁴⁷⁰

Additionally, the lack of legal clarity regarding liability in computer code errors is a problem, as it is unclear who should be held responsible - the programmer, the client, or the platform. Liability in open source models is not inherent, but limitations of liability are often expected. Public blockchain software is typically issued under open source software licenses, which disclaim liability for any claims arising from the software. However, the liability of software developers of public blockchain systems has been examined and the "veil of decentralization" is often used to shield non-compliant systems from liability. Effective enforcement of existing norms is important to counteract this.⁴⁷¹

There is also a question what type of liability standards shall be applicable, and whether it should be, e.g. strict liability for any fault in a product/for negligence or a reasonable-care standard.⁴⁷² Moreover, it should be also considered whether and to what extent classic national measures tackling liability issues (e.g. present in criminal law) could be applicable.

*Depending on the specific area at stake, access points for attributing liability (such as in the form of a damage claim) and compliance may be the same or different actors, depending on the specific circumstances at stake and area of the law concerned. It must be stressed that even where no suitable governance arrangements exist, law enforcement agencies may bring claims against actors that participate in the administration of the given blockchain usecase, such as for instance miners, nodes, core developers, Internet Service Providers or even users.*⁴⁷³

DATA LOCALISATION REQUIREMENTS

Data localisation requirements and data retention rules are significant for blockchain technology since it is a type of database, and all EU data regulations must be followed. Compliance with data localisation requirements may be challenging for transnational data networks, but the EU's new regulation on the Free Flow of Non-Personal Data ensures the free movement of data in the internal market, including Distributed Ledger Technology (DLT).

BLOCKCHAIN, SSI AND EIDAS

⁴⁶⁸ [*Blockchain \(cnil.fr\)](https://www.cnil.fr/actualites/nouvelles/nouvelle-4777.aspx), p. 3

⁴⁶⁹ Michèle Finck, *Blockchain Regulation and Governance in Europe*, Cambridge University Press, 2019, Chapter 2.

⁴⁷⁰ Study on Blockchains Legal, governance and interoperability aspects (SMART 2018/0038), p. 46

⁴⁷¹ Ibid, p. 47

⁴⁷² Ibid.

⁴⁷³ Primavera De Filippi, Aaron Wright, 'Blockchain and The Law: The Rule of Code', Harvard University Press (March 2019), available at <https://www.hup.harvard.edu/catalog.php?isbn=9780674241596&content=toc>, Chapter V.

Self-Sovereign Identity (SSI) systems, designed according to privacy-by-design principles, can be a blockchain-based solution used to address inefficiencies in traditional identity verification processes. SSI systems allow users to disclose identity information in a selective manner. Moreover, verifiable credentials issued by such systems are revocable. Nevertheless, there are challenges in uniting eIDAS with SSI, as eIDAS has been developed taking into account hierarchical trust centers and trust schemas.

Furthermore, traditional eIDAS systems are linked with blockchain and smart contracts. The examples are “message signatures” and “signed” transactions, which are the outcome of public-key based cryptography, same as HTTPS or qualified digital certificates issued by qualified trust service providers.

Public blockchains are permissionless, verifying only the authentication and integrity of messages, while eIDAS aims to ensure credentials' authenticity, integrity, and validity, by ensuring no tampering of the credentials, they are within their validity period, the issuer did not revoke them, they have intended semantics and are provided by an issuer who is trustworthy. At the protocol level, blockchain systems such as Ethereum only verify the authenticity and integrity of the message signed, without attempting to validate the trustworthiness of the actor signing a transaction. These systems solely check if the message meets the criteria for a valid signature.

INTELLECTUAL PROPERTY LAW FRAMEWORK

TRUSTEE envisages to facilitate the mediation of access to different datasets provided by various parties for secondary use purposes as well as AI/ML models for computations to be performed on TRUSTEE. In the above chapters, mainly the classification between personal and non-personal data has been made as to the type of data that could potentially be accessed to via TRUSTEE. However, another classification should also be made within the legal framework, which pertains to intellectual property (“IP”) rights not only over datasets to which access will be mediated but also AI/ML models to be uploaded to and used via TRUSTEE.

There might be different types of intellectual property rights protecting certain types of data/datasets and AI/ML models, which the data and model providers, both representing “users” of TRUSTEE and (potentially) types of “rightholders”, would like to grant access to consumers of TRUSTEE for certain actions to be conducted on TRUSTEE. Not every data, dataset or machine learning model are protected by intellectual property rights. The protection of datasets and AI/ML models is contested in literature. Nonetheless, the application of the protection from intellectual property rights is subject to the satisfaction of the legal requirements, which will be elaborated on below. Where protection is granted, the holders of intellectual property rights can exercise certain exclusive rights over the protected subject matter. Therefore, it is important to assess whether certain intellectual property rights (and limitations flowing thereof) need to be considered for the arrangement and facilitation of access to datasets and models on TRUSTEE. The below sub-chapters will refer to and analyze, in particular, copyright law and trade

secret law frameworks stemming from the Berne Convention⁴⁷⁴, the Infosoc Directive⁴⁷⁵, Computer Programs Directive⁴⁷⁶, Database Directive⁴⁷⁷, CDSM Directive⁴⁷⁸, and the Trade Secret Directive⁴⁷⁹.

INTRODUCTION TO THE EUROPEAN INTELLECTUAL PROPERTY LAW FRAMEWORK

EUROPEAN COPYRIGHT LAW FRAMEWORK – THE BASICS

Within the European Union, copyright law framework is primarily based on directives, which aim for only a limited level of harmonization. This leads to a considerable level of fragmentation among Member States' laws. This should be kept in mind and the national laws of each Member State should be consulted in order to understand specific applications of certain provisions to be mentioned below. It is up to the Data Provider and Model Provider and to assess the limitations stemming from the national legal acts and to indicate that compliance with them is one of the conditions to be allowed to use the datasets and/or models. Regardless of this, there are two common requirements that need to be satisfied for copyright protection - European copyright law framework protects:

- literary, artistic and scientific works⁴⁸⁰,
- which are original.

First criterion sets the relevant subject matter for copyright protection in a non-exhaustive manner. Article 2(1) of Berne Convention keeps the scope of the subject matter broad, covering “every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression”⁴⁸¹. Some examples include books, illustrations as well as software/computer programs⁴⁸² and source code (“The expression of a program in a programming language is known as the source code of the program.”⁴⁸³), but may even stretch to daily engagements, such as online comments or reviews. It is up to the countries which are part of the Convention to decide whether for copyright protection to be granted, the work has to be fixed in a material form or not.⁴⁸⁴

Second criterion, however, is harder to satisfy than the first one and has been developed primarily by the jurisprudence of the Court of Justice of the European Union in the EU. For a work to be deemed as original, it needs

⁴⁷⁴ Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979) [Online]. Available at: <https://www.wipo.int/wipolex/en/treaties/textdetails/12214>.

⁴⁷⁵ Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society [Online]. Available at: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32001L0029>.

⁴⁷⁶ Directive 2009/24/EC of 23 April 2009 on the legal protection of computer programs (‘Software Directive’) [Online]. Available at: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32009L0024>.

⁴⁷⁷ Directive 96/9C of 11 March 1996 on the Legal Protection of Databases [Online]. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31996L0009>.

⁴⁷⁸ Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (‘CDSM Directive’) [Online]. Available at: <https://eur-lex.europa.eu/eli/dir/2019/790/oj>.

⁴⁷⁹ Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [online]. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L0943>.

⁴⁸⁰ Berne Convention, Article 2(1).

⁴⁸¹ Berne Convention, Article 2(1).

⁴⁸² TRIPS Agreement Article 10(1) states that “computer programs, whether expressed in source code or in object code, will be protected as literary works pursuant to the Berne Convention.”

⁴⁸³ The Software Interface Between Copyright and Competition Law: A Legal Analysis of Interoperability in Computer Programs, Ashwin van Rooijen Wolters Kluwer [2010] 2010 12.

⁴⁸⁴ Berne Convention, Article 2(2).

to be “the author’s own intellectual creation”.⁴⁸⁵ This requires the author of the work to be able to make “free and creative choices”⁴⁸⁶ along with having his “personal touch”⁴⁸⁷ in the work. Even extracts from a newspaper article comprising of 11 words may be protected by copyright “if that extract contains an element of the work which, as such, expresses the author’s own intellectual creation.”⁴⁸⁸ However, where “the realisation of a subject matter has been dictated by technical considerations, rules or other constraints, which have left no room for creative freedom”⁴⁸⁹, the originality requirement will not be satisfied.

Ideas as such are not protected by copyright but rather the expression of ideas in a manner representative of the author’s intellectual creation for copyright protection. This is called idea/expression dichotomy in copyright law. Therefore, any facts, statement of facts or data, such as non-personal data but also personal data on name, age or occupation, cannot be protected by copyright but reside in the public domain free to be used by anyone (subject to limitations from other legal acts, such as data protection laws).⁴⁹⁰

The exclusive/economic rights conferred by copyright last till the end of 70 years after the death of the author of the work.⁴⁹¹ Such rights grant the copyright holder privilege to prohibit other parties from conducting certain acts with the copyright holder’s work. They mainly include the exclusive rights to:

- (1) “authorise or prohibit direct or indirect, temporary or permanent reproduction by any means and in any form, in whole or in part” of the work⁴⁹²;
- (2) “authorise or prohibit any communication to the public of their works, by wire or wireless means, including the making available to the public of their works in such a way that members of the public may access them from a place and at a time individually chosen by them”⁴⁹³;
- (3) “authorise or prohibit any form of distribution to the public by sale or otherwise” of both the original work or copies thereof.⁴⁹⁴ Conducting these acts without the permission of the rightholder leads to the infringement of the copyright in the work.

Moral rights, on the other hand, protect the attachment of the work to its author. They concern rights such as the right to be recognized as the author of the work or right to protect the integrity of the work. The regulation of moral rights, such as the rules on their duration and transferability, are subject to the laws of each member state and are not harmonized on an EU level.⁴⁹⁵

In order to use a work protected by copyright, one needs to obtain permission from the copyright owner. This is possible either by obtaining a license through which the rightholder gives permission for certain uses of his work

⁴⁸⁵ *Infopaq International A/S v Danske Dagblades Forening* (C-5/08) [2009] ECLI:EU:C:2009:465, para 37, 45.

⁴⁸⁶ *Football Dataco Ltd, Football Association Premier League Ltd, Football League Limited, Scottish Premier League Ltd, Scottish Football League and PA Sport UK Ltd v Yahoo! UK Limited, Stan James (Abingdon) Limited, Stan James PLC and Enetpulse APS*, (C-604/10) [2012] ECLI:EU:C:2012:115, para 38.

⁴⁸⁷ *Eva-Maria Painer v Standard VerlagsGmbH, Axel Springer AG, Süddeutsche Zeitung GmbH, SPIEGEL-Verlag Rudolf AUGSTEIN GmbH & Co KG and Verlag M. DuMont Schauberg Expedition der Kölnischen Zeitung GmbH & Co KG* 8 (C-145/10) [2011] ECLI:EU:C:2011:798, para 92.

⁴⁸⁸ *Infopaq* (n 305), para 48.

⁴⁸⁹ *Cofemel — Sociedade de Vestuário SA v G-Star Raw CV* (C-683/17) [2019] ECLI:EU:C:2019:721, para 31.

⁴⁹⁰ EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility, Inga Graef, Wolters Kluwer [2016] p.134.

⁴⁹¹ Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights, Article 1(1) [online]. Available at: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32006L0116>.

⁴⁹² Infosoc Directive, Article 2.

⁴⁹³ Infosoc Directive, Article 3.

⁴⁹⁴ Infosoc Directive, Article 4.

⁴⁹⁵ Berne Convention, Article 6^{bis}(2); Infosoc Directive, Recital 19.

while still enjoying copyright ownership or by transferring either all or certain exclusive rights granted by copyright to others. However, one does not need to obtain permission from the copyright owner for all types of uses. EU copyright law provides various opportunities for uses without permission:

- (1) Using a trivial or insignificant part of a copyrighted work (“de minimis use”) might be possible without the need to obtain the permission of the rightholder. However, this is often difficult to argue as even short snippets or small parts of a work might be reaching the threshold of originality and thus are still protected by copyright.
- (2) When a copyrighted work is made available under an open license, the user may not need to obtain permission for re-using the work, as they are informed in advance via such license that certain uses of the copyrighted work are already allowed by the rightholder. An example of open licenses is Creative Commons Licenses.⁴⁹⁶
- (3) Copyright law limits the exclusive rights of the rightholders by allowing for certain exceptions to their exclusive rights granted by copyright, which are mainly optional to transpose and therefore, each Member State can decide which exception they would like to implement to their national laws. Copyright exceptions may be working differently from one another. Some allow only for certain uses (i.e. reproduction of the work but no distribution of it). Some allow certain uses only for specific categories of re-users (i.e. researches, cultural heritage institutions, press publishers). Some allow re-uses for specific purposes (i.e. research and teaching, criticism or review, text and data mining) and these might be subject to certain conditions (i.e. having lawful access to the copyrighted work from the beginning, mentioning the name of the work’s author properly).

Within the initial legal analysis presented in this deliverable, the text and data mining exception, which was presented as an exception mandatory to be implemented by all member states in the CDSM Directive⁴⁹⁷, will be reviewed as the exception that might have the most relevance to TRUSTEE. Where relevant, other exceptions and limitations EU copyright law foresees and which might have relevance to TRUSTEE will be analysed and considered in the following iterations of this deliverable.

TRUSTEE aims to accommodate data mining functionalities and queries to search the TRUSTEE knowledge repository in order to locate data a TRUSTEE consumer needs. Offering such functionality requires the consideration of the text and data mining (“TDM”) exceptions that allow for acts of reproduction and extraction of copyright-protected or database-protected subject matter, which the party conducting the text and data mining has lawful access to. The two TDM exceptions are distinct from each other, as the parties allowed to conduct the TDM activities and the purposes of the TDM activities allowed differ.

Article 3 of the CDSM Directive allows reproductions and extractions of the copyright or database-protected subject matter concerned to “research organisations and cultural heritage institutions”, where the TDM is conducted “for the purposes of scientific research”.⁴⁹⁸ The exception requires the secure storage of copies made while conducting text and data mining activities. Article 4 of the same Directive extends the text and data mining exception beyond the research organizations and cultural heritage institutions and allows “reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining” by anybody. However, this limitation applies only if the rightholder of the subject matter to be used for text and data mining has not expressly forbid the use of the subject matter in an appropriate manner (i.e. by machine-readable means where the content has been made publicly available online).

These exceptions will not be applicable where the subject matter reproduced or extracted for text and data mining purposes via TRUSTEE is not protected by copyright or database rights. Nevertheless, even when such subject matter exists among the material used for text and data mining via TRUSTEE, data provider will be granting access

⁴⁹⁶ Creative Commons Licenses [online]. Available at: <https://creativecommons.org/licenses/>.

⁴⁹⁷ CDSM Directive, Articles 3 & 4.

⁴⁹⁸ CDSM Directive, Article 3.

to their datasets for certain uses, where such permission could also cover text and data mining. Such arrangement could be made part of the “Security & Trust Manager” module. Further analysis, when necessary, will follow.

Protection of computer programs

Copyright protection also extends to computer programs, the protection of which is regulated under the Directive on the legal protection of computer programs (‘Software Directive’). The Directive itself does not give a definition of computer programs, but these can be described as a “set of instructions aimed at having computer hardware perform a specific task”. Computer programs, more specifically all forms of the expression of computer programs, which also extend to the source code and object code of a computer program, are protected by copyright as literary works within the meaning of the Berne Convention, also including the preparatory design materials of the computer program. The originality requirement is equally applicable here, meaning a work, in this case the computer program, needs to be original in the sense of carrying the personal stamp of its author. Ideas and principles underlying any part of the computer program, thus the logic, algorithms and programming languages to the extent that they comprise ideas and principles, however, are not protected by copyright. This is because copyright protects the expression of ideas and principles underlying the software (i.e. “arrangement of the instructions that configure it, its sequence, structure, and the relationship between the parts”), not the ideas and principles themselves. Based on the idea-expression dichotomy to the protection of computer programs, data files as well may not qualify for the copyright protection granted under the Software Directive as these are also not accepted as forms of expression of the computer program.

The primary reason of covering the protection of computer programs from the perspective of copyright law is the exclusion of protection of software via patents. This is excluded in Europe by law under article 52(2) of the European Patent Convention, despite the highly debated nature of the topic in Europe in practice where the European Patent Office itself at times argues that the law does not prevent the patentability of software. Though a computer program as such is not patentable, a patent could be granted to a computer-implemented invention where a technical problem is solved in a novel and non-obvious manner. Where necessary, further iterations of the legal framework will cover the analysis of patent protection of computer programs.

The author of a computer program, according to the Directive, is primarily “the natural person or group of natural persons who has created the program”. It might also be that someone else is designated as the right holder, if the legislation of a member state allows it and designates a legal person as such. If a computer program is created by an employee while executing his duties or following the instructions of his employer, the employer is to have the right to exercise the exclusive economic rights borne out of the copyright in the computer program, unless the contract between the employer and the employee states otherwise.

The copyright in a computer program grants the right holder the right to authorise the following:

- “the permanent or temporary reproduction of a computer program by any means and in any form, in part or in whole; in so far as loading, displaying, running, transmission or storage of the computer program necessitate such reproduction;
- the translation, adaptation, arrangement and any other alteration of a computer program and the reproduction of the results thereof, without prejudice to the rights of the person who alters the program;
- any form of distribution to the public...”

There are certain exceptions to the above mentioned exclusive rights. A primary exception that could be relevant for the purposes of TRUSTEE is that the person who has a right to use a copy of a computer program does not require the authorisation of the rightholder to “observe, study or test the functioning of the program in order to determine the ideas and principles which underlie any element of the program if he does so while performing any of the acts of loading, displaying, running, transmitting or storing the program which he is entitled to do.” Such acts shall not be considered infringing the copyright in the computer program and the parties cannot decide on any contractual provisions contrary to this exception, which would otherwise be considered null and void. This

exception would apply when and to the extent that TRUSTEE will offer the possibility to offer access to different source code provided by TRUSTEE users to perform certain computations with. This can especially be relevant with respect to the TRUSTEE software development kit aimed to be provided as a toolset to TRUSTEE users, where the toolkit will present choices of code samples to those wanting to integrate TRUSTEE framework to their own applications. It is advised to outline allowed uses of such sample codes (depending on whether they are protected by copyright to begin with or not) under the terms of use of TRUSTEE, where the above mentioned exception to the copyright in the computer program should be taken into account.

Protection of databases

Directive (EC) 96/9C of 11 March 1996 on the Legal Protection of Databases ('Database Directive')⁴⁹⁹ was introduced to provide a certain level of harmonisation for the protection of databases in the internal market. The Directive defines a database as “a collection of independent works, data or other materials, which are

- arranged in a systematic or methodical way and,
- individually accessible by electronic or other means.”⁵⁰⁰

Additionally, the author of a database is designated as the natural persons creating the database itself or those who are assigned as the right holders by a national legislation.⁵⁰¹

Full copyright protection is granted to the expression of databases by the Directive, if the selection or arrangement of the contents of the database results from the author's own intellectual creation⁵⁰², whereby the author “expresses his creative ability in an original manner by making free and creative choices.”⁵⁰³ This means that only the structure of a database which is original can be protected by copyright, regardless of the copyright status of the contents of the database. If the structure of the database results from “technical considerations, rules or constraints which leave no room for creative freedom”⁵⁰⁴, the copyright protection to the database will not apply. This is regardless of the fact that the author of the database applied a significant level of labour and skill in creating the data making up the database or to selecting or arranging it, unless such labour and skill express originality in the selection or arrangement of data.⁵⁰⁵ It is argued however that most modern electronic databases lack the sufficient originality in their arrangements.⁵⁰⁶ This might also be the case in the collection of raw data for research purposes⁵⁰⁷ or where the database is automatically created and the data within is selected by a software, without the involvement of a human author. However, some also argue that “in the context of AI and big data the selection and combination of training data, structuring and weighing of [...] systematic structures which form basis for generating inferred data could

⁴⁹⁹ Database Directive.

⁵⁰⁰ Database Directive, Article 1(2).

⁵⁰¹ Database Directive, Article 4.

⁵⁰² Database Directive, Article 3(1).

⁵⁰³ Football Dataco Ltd, Football Association Premier League Ltd, Football League Ltd, Scottish Premier League Ltd, Scottish Football League, PA Sport UK Ltd v Yahoo! UK Ltd, Stan James (Abingdon) Ltd, Stan James plc, Enetpulse ApS (C-604/10) [2012] ECLI:EU:C:2012:115, para 38.

⁵⁰⁴ Ibid, para 39.

⁵⁰⁵ Ibid, para 42.

⁵⁰⁶ Legal study on ownership and access to data, Final report – Study, European Commission 2016, p. 13 [online]. Available at: <https://op.europa.eu/s/n2Qc>. See also Bently, L., Derclaye, E. and others, Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases, p. 26: “The creation of database content has evolved in the last decades from the manual gathering of existing data, over automatic processes of data collection, even to automatic creation of data (e.g. sensor-generated data).”

⁵⁰⁷ Linda Kuschel and Jasmin Dolling, Access to Research Data and EU Copyright Law 13 (2022) JIPITEC 247. The authors argue in para 11 that collection of raw data in the context of research “will generally be comprehensive in nature and therefore not subject to an individual selection; the arrangement in turn will follow logical criteria (eg, time, quantity, size), as the representation should meet scientific standards and be as clear and verifiable as possible. Therefore, there is little room for creative selection or arrangement.”

generally qualify as ‘creative’ compilation of independent elements. However, due to the rather high requirements for an own intellectual creation which also have to be applied to database works only an outstanding selection and combination of data (e.g. specific training datasets) could be protected under copyright law.”⁵⁰⁸

Copyright protection of the expression of databases grant the author of a database exclusive rights to carry out or to authorize:

- “temporary or permanent reproduction by any means and in any form, in whole or in part;
- translation, adaptation, arrangement and any other alteration;
- any form of distribution to the public of the database or of copies thereof,
- any reproduction, distribution, communication, display or performance to the public of the results of the acts”⁵⁰⁹ from its translation, adaptation, arrangement or any other alteration.

However, where the performance of the above mentioned acts by a lawful user of a database or a copy of it is necessary to access the contents of the databases and for the normal use of the contents, the lawful user does not require the permission of the author of the database.⁵¹⁰

The above mentioned copyright protection does not extend to the contents (data) of a database.⁵¹¹ Nonetheless, the Database Directive provides a dual regime of protection, where the contents of a database can be protected by what is called a *sui generis* right, even though the structure of the database cannot be protected by copyright due to not reaching the required threshold of originality or the contents of the database themselves are not eligible for copyright protection.⁵¹² A database may qualify for protection under the *sui generis* right, if:

- the maker of the database can show that they have made “qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents”⁵¹³,
- the maker of the database is an EU citizen or a resident in the EU.⁵¹⁴

Recital 41 of the Database Directive also recognizes the maker of a database as “the person who takes the initiative and risk of investing” in the obtaining, verification or presentation of the contents of the database. In situations, especially considering the current data sharing networks and platforms, which are “characterised by cooperative structures involving a multitude of players [...] each participant bears at least the risk for its own investment and can, therefore, qualify as database maker leading to joint ownership of the database.”⁵¹⁵

⁵⁰⁸ Matthias Leistner and Lucie Antoine, IPR and the use of open data and data sharing initiatives by public and private actors (2022) p. 49 [online]. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf).

⁵⁰⁹ Database Directive, Article 5.

⁵¹⁰ Database Directive, Article 6(1).

⁵¹¹ Database Directive, Article 3(2).

⁵¹² Database Directive, Chapter III. See also Graef, p.135.

⁵¹³ Database Directive, Article 7(1).

⁵¹⁴ Database Directive, Article 11(1).

⁵¹⁵ Matthias Leistner and Lucie Antoine, IPR and the use of open data and data sharing initiatives by public and private actors (2022) [online] available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf), p. 55: “in the context of access rights, joint ownership leads to legal uncertainty and high information and transaction costs, e.g. for identifying all rightholders and carrying out negotiations. However, for the time being, it would not be reasonable to allocate the *sui generis* right differently as potential data sharing scenarios are very diverse and data markets are still developing dynamically. Consequently, it should be left open to the participants of data sharing networks to agree on a particular allocation of the *sui generis* right by means of contract.”

The *sui generis* right can be transferred, assigned or granted to other parties by way of a contractual license.⁵¹⁶ Different from the term of protection of copyright, the *sui generis* protection lasts for fifteen years from the first of January of the year following the date of completion of making the database.⁵¹⁷

It might be that a database meets the requirements of both copyright protection, for its structure, and also the *sui generis* protection, for its contents. However, these rights are protecting two different aspects, structure and contents, of the same object, which is the database. The author of a database will have the exclusive right to “prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database.”⁵¹⁸ The rights to prevent the extraction and/or re-utilization mirror the rights of reproduction and communication to the public, which the authors of original databases have per Article 5 of the Database Directive. Despite these exclusive rights of the maker of a database, the maker cannot prevent a lawful user of the database from extracting and/or re-utilizing insubstantial parts of its contents, if the database was made available to the public.⁵¹⁹ A lawful user of a database or a copy of it does not also need the authorisation of the author of the database to conduct certain acts, such as translation or adaptation⁵²⁰, that are necessary to access the contents of the database and use the contents of the database in a normal manner.⁵²¹ Nevertheless, a lawful user of a database, which was made available to the public, “may not perform acts which conflict with normal exploitation of the database”.⁵²² National laws of Member States may provide for exceptions to the *sui generis* database right, allowing a lawful user of a database made available to the public to extract or re-utilize a substantial part of the contents of the database for certain purposes without the permission of the holder of the database right, such as for teaching or scientific research.⁵²³ Finally, it is important to note that the Directive renders any contractual arrangements contrary to the rights and obligations of lawful as described in the Directive as null and void.⁵²⁴

TRUSTEE aims to facilitate the sharing of open and FAIR⁵²⁵ databases and datasets. The tools to be provided on TRUSTEE and actions envisaged to be done with the datasets most likely will require the extraction of either all or

⁵¹⁶ Database Directive, Article 7(3).

⁵¹⁷ Database Directive, Article 10(1).

⁵¹⁸ Database Directive (n 298), Articles 7(1), 7(2). Extraction refers to the permanent or temporary transfer of the whole or a substantial part of the contents of the database to another medium, by any means or in any form. Re-utilizing refers to any form of making available to the public the whole or a substantial part of the contents of the database by distributing or renting copies, through on-line or other forms of transmission.

⁵¹⁹ Database Directive (n 298), Article 8(1). Other rights and obligations of lawful users also apply. See also Legal study on ownership and access to data, Final report – Study, European Commission 2016 [online]. Available at: <https://op.europa.eu/s/n2Qc>, p. 15: “This appears to make licensing under database right a ‘once and for all’ arrangement rather than being able to be restricted to specific fields of use. It is not clear for what purpose the framers of the legislation included this provision but it could hamper the exploitation of databases which, for any reason are made available to the public either in the course of commercial licensing or otherwise, through licensing since such a loss of control may incentivise database right holders from making licences available on any terms.”

⁵²⁰ Further list of these acts are given under Article 5 of the Database Directive.

⁵²¹ Database Directive, Article 6(1).

⁵²² Database Directive, Article 8(2).

⁵²³ Database Directive, Article 9.

⁵²⁴ Database Directive, Article 15. This is, however, in the case where a database is protected either by copyright or by the *sui generis* right under the Directive. If not, the Directive “does not prevent the adoption of contractual clauses concerning the conditions of use of such a database.” See *Ryanair Ltd v PR Aviation BV* (C-30/14) [2015] ECLI:EU:C:2015:10, para. 39.

⁵²⁵ Javier de la Cueva and Eva Méndez, “European Commission Report of Study – Open Science and Intellectual Property Rights: How can they better interact? State of the art and reflections” (2022) p. 28 [online]. Available at: https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/open-science-and-intellectual-property-rights_en

substantive parts of a database.⁵²⁶ Where the holder of a database sharing access to it via TRUSTEE permits the extraction or re-utilization of substantial parts of a database for certain actions, the re-use can function smoother as the database maker assures not to assert his exclusive rights in the database. TRUSTEE can aim to encourage database and dataset holders to provide access to their datasets with as many open terms as possible. With the Security and Trust manager tool and the contractual level aimed to be developed, TRUSTEE can contribute to the discussions on best practices for providing access to datasets.

TRADE SECRETS / KNOW-HOW

There might be important information developed through the course of research and innovation in trade, which the parties working on its creation and progression may not be able to protect with the traditional IP rights (such as copyright as mainly discussed above) or may not find it convenient to invest in IP protection. Regardless of the obstacles to considering the protection from the traditional IP rights' side, keeping the information/know-how that is valuable to the business confidential could also attract a certain level of protection as "trade secrets". Please note that trade secret protection is mainly not considered as part of the intellectual property realm.

In the EU, the protection of trade secrets against unlawful acquisition, their use and disclosure are harmonized to a certain extent by the Trade Secret Directive (also called Know-How Directive). The aim of the protection of trade secrets is to maintain and foster the incentive to innovate on part of businesses by ensuring that, through trade secret protection, the time and investment businesses devote to innovations as well as the competitive advantage they gain by keeping certain information confidential are safeguarded. For the information to be protected as a trade secret, the individual pieces of information do not need to be novel like the requirement under patent protection, nor they need to be registered or protected by copyright. The protection does not differentiate between different types of information, where the main criteria for protection is the confidential nature of the information, and the protection continues as long as the information is maintained as a secret.

The Directive sets the following requirements for information to be regarded as a trade secret:

- "secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question;
- it has commercial value because it is secret;
- it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret."

The definition covers "know-how, business information and technological information where there is both a legitimate interest in keeping them confidential and a legitimate expectation that such confidentiality will be preserved." Examples of trade secrets may include R&D data, computer programs, computer databases, information regarding product technology, and as often observed, AI/ML systems. Additionally, the commercial value of the trade secret as required in the definition can either be actual or potential. Furthermore, some examples of reasonable steps to keep the information confidential could be storing the information in a safe medium or signing non-disclosure agreements with any parties that have access to the information.

The Directive is clear in its emphasis on not creating any exclusive right to know-how or information protected as trade secrets. This means that, unlike intellectual property rights described above, the trade secret holder will not have an exclusive right to their use, which he can enforce against others. Thus, any independent discovery or

⁵²⁶ Matthias Leistner, "Big Data and the EU Database Directive 96/9/EC: Current Law and Potential Reform" in S Lohsse, R Schulze and D Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Hart Publishing 2018) 41.

creation of the same know-how is to be regarded as lawful, as well as the reverse engineering of a product (“observation, study, disassembly or testing” of it) that is either in public domain or has been lawfully acquired, unless the latter is contractually limited. If the trade secret was obtained through lawful means as described above, such circumstance does not warrant a trade secret protection.

IPRS AND DATA/DATASETS

This sub-chapter will evaluate the possibility of exerting rights and control over data via certain intellectual property rights. It will not engage in an extensive debate on the concept of ownership in data but rather analyze which rights might be assigned to data by way of the existing intellectual property rights. As a note however, it is important to state that rather than the intellectual property rights mentioned below, contractual means for the assignment of rights to data might be a preferred route, which is also favored in the recent Data Act Proposal.⁵²⁷ This will be further mentioned in the chapter below on contract law and licensing considerations.

COPYRIGHT PROTECTION

Under chapter 1.10.1 outlining the basics of European copyright law, the idea-expression dichotomy is emphasized, which declares that only the expression of ideas, rather than the ideas themselves, are protected by copyright. Data as “‘raw’ knowledge and information” is deemed to be categorized as ideas rather than expressions.⁵²⁸ Stemming from this, any facts, statement of facts or data, such as non-personal data but also personal data on name, age or occupation, cannot be protected by copyright but reside in the public domain free to be used by anyone (subject to limitations from other legal acts, such as data protection laws).⁵²⁹

There is extensive debate in academic circles on the extent of application of intellectual property rights to digital data in the traditional sense and ownership of data (including whether such concept exists). For works subject to copyright protection that can also be rendered as digital data, measures with respect to ownership and exclusive rights, as outlined in previous sub-chapters, are generally more straightforward as the use of such works by third parties usually is regulated under licensing agreements where the parties determine the terms of use of the work concerned. When it comes to non-original works that are rendered as digital data, however, as these will not be subject to protection of copyright as the level of originality is not met in the underlying work, it is argued that licensing agreements would not be applicable here, though any contractual limitations as to the terms of use might still be valid.⁵³⁰ This will be elaborated upon under the sub-chapter on licensing arrangements.

⁵²⁷ Please see Chapter 1.4.3 for detailed information on the Data Act proposal. Please also see Margoni, Thomas, Ducuing, Charlotte, & Schirru, Luca. (2023). Data property, data governance and Common European Data Spaces. <https://doi.org/10.5281/zenodo.7867881>, p. 11: “The Data Act proposal dedicates significant attention to contracts and, interestingly, to contract regulation, especially in cases of B2B data transactions. Due to the limited availability of (intellectual) property rights on data, contract regulation is leveraged by the proposal as a significant means to allocate data rights indirectly.”

⁵²⁸ Ivan Stepanov (2020) Introducing a property right over data in the EU: the data producer’s right – an evaluation, *International Review of Law, Computers & Technology*, 34:1, 65-86, DOI: [10.1080/13600869.2019.1631621](https://doi.org/10.1080/13600869.2019.1631621), p. 71.

⁵²⁹ EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility, Inga Graef, Wolters Kluwer [2016] p.134.

⁵³⁰ Data sharing as a service: Will data services remove intellectual property rights from the picture, and at what cost? Timelex, European Commission [online]. Available at: https://data.europa.eu/sites/default/files/report/intellectual_property_rights_report.pdf, p. 6-8. The study recommends the framing of terms of use of non-original data within a “subscription agreement”: “when digital data is made available purely on the basis of a subscription agreement, the data holder is again free to impose any

DATABASE *SUI GENERIS* PROTECTION

The copyright protection granted to the expression of a database, where its structure is original in the sense that the selection or arrangement of its contents (compilations of data) result from the database maker's own intellectual creation, does not extend to the contents (data) of a database.⁵³¹ This also applies to *sui generis* database right, where the right is granted only to "data structured in a 'database'".⁵³² Nevertheless, database *sui generis* protection is still considered to "play a crucial role for the European data economy and for the envisaged objective of fostering data sharing, access and re-use as its scope of protection may potentially cover many typical big data and AI scenarios."⁵³³

The protection of the contents of a database is debated more especially for big data or AI-related data due to "the assumption that in most big data situations the crucial condition of a 'substantial investment' will not be fulfilled."⁵³⁴ This is due to the requirement under the Database Directive that in order to gain the *sui generis* protection, a substantial investment in obtaining, verifying or presenting the contents of a database needs to be made. The investment put into creating the materials which constitute the contents of a database does not qualify for the *sui generis* protection, meaning individual data making up the content of a database is not protected by such database right⁵³⁵, despite there being a substantial investment made for its creation. This is also regardless of the increase in the value of the database due to the investment put into making up the contents of the database. Reasoning behind such analysis, stemming from the jurisprudence of the CJEU, is the broad scope of the *sui generis* database right where the right can act as an "exclusive property right against the use of data", which is "solely available from one

constraints, including the non-transferability of the data itself, based on the general principle of contractual freedom." As an example case, see *Ryanair Ltd v PR Aviation BV* (C-30/14) [2015] ECLI:EU:C:2015:10.

⁵³¹ Database Directive (n 298), Article 3(2) and Recital 45: "Whereas the right to prevent unauthorized extraction and/or re-utilization does not in any way constitute an extension of copyright protection to mere facts or data." See also Legal study on ownership and access to data, Final report – Study, European Commission 2016 [online]. Available at: <https://op.europa.eu/s/n2Qc>, p. 13: "Where copyright does subsist, this confers a property right upon the holder of the database so that the copyright can be assigned, mortgaged or licensed like any other form of property. It also enables the owner to prevent third parties from copying the data without permission, regardless of whether the data is secret." See also Francesco Banterle, Data ownership in the data economy: a European dilemma, (August 1, 2018). EU Internet Law in the digital era, (edited volume based on the REDA 2017 conference), Springer, (2018), available at SSRN: <https://ssrn.com/abstract=3277330>, p. 5: "copyright applies to the specific form in which the compilation of data is structured (Database Directive Recital 15), and not to the underlying raw data. The same data can be taken from the original selection and reshuffled without reproducing the selection/arrangement."

⁵³² P. Bernt Hugenoltz, 'Data Property in the System of Intellectual Property Law: Welcome Guest or Misfit?' in S Lohsse, R Schulze and D Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Hart Publishing 2018) 89. See also Francesco Banterle, Data ownership in the data economy: a European dilemma, (August 1, 2018). EU Internet Law in the digital era, (edited volume based on the REDA 2017 conference), Springer, (2018), available at SSRN: <https://ssrn.com/abstract=3277330>, p. 6: "The Database Directive did not intend to create a new right in data (Recital 46 of the Database Directive), nor to protect mere facts or data (Recital 45), but rather to keep the semantic information in the public domain."

⁵³³ Matthias Leistner and Lucie Antoine, IPR and the use of open data and data sharing initiatives by public and private actors (2022) p. 49 [online]. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf).

⁵³⁴ Matthias Leistner, 'Big Data and the EU Database Directive 96/9/EC: Current Law and Potential Reform' in S Lohsse, R Schulze and D Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Hart Publishing 2018) 27.

⁵³⁵ Graef, p. 136. See also *Fixtures Marketing Ltd v. Oy Veikkaus Ab* (C-46/02) [2004] ECLI:EU:C:2004:694, para. 39.

particular source [...] and, thus, cannot be obtained with comparable investments otherwise.”⁵³⁶ In these cases, it is stated that “the *sui generis* right factually would amount to a protection of the generated data, going way beyond its objective to protect the database maker’s investment.”⁵³⁷

However, the above discussion on the activities to which the substantial investment criteria apply becomes especially tricky for the situations where there is a difficulty in drawing “a clear line between the *collection* and the mere *creation* of data”.⁵³⁸ This is pointed out to be the case particularly for the “‘producers’ of sensor or machine-generated data” and databases, whose investments are observed to be made to the creation of data as a by-product of their main activities rather than to the other activities mentioned above, and therefore, such investments would not be within the scope of the *sui generis* protection.⁵³⁹

Therefore, if a substantial investment was made only to the creation of the data constituting the database, this does not warrant any *sui generis* protection whereas “the resources used to search for existing materials and to collect them in the database will count as a ‘substantial investment’”⁵⁴⁰ and can warrant *sui generis* protection. Such protection would grant the database maker the right to object to extraction and re-utilization of substantial parts of the contents of his database/dataset.⁵⁴¹ Yet again, these exclusive rights do not bring with them rights to the individual data making up the database.⁵⁴²

The interpretation of the substantial investment criteria of the *sui generis* database right has triggered plenty of discussions on whether the *sui generis* database right is capable of fitting to the current realities of the data economy. Currently, most critics argue that the *sui generis* right could “intensify the problem of de facto control over data leading to significant access problems and even hold-up potential” due to this “legal uncertainty in regard to core elements of *sui generis* protection, such as the distinction between collection and creation of data”.⁵⁴³ Though no updates to the Directive or to the right itself are on the horizon, any relevant changes or discussions will be reflected in the upcoming iterations of this deliverable.

⁵³⁶ Matthias Leistner and Lucie Antoine, IPR and the use of open data and data sharing initiatives by public and private actors (2022) p. 50 [online]. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf).

⁵³⁷ Ibid.

⁵³⁸ Matthias Leistner and Lucie Antoine, IPR and the use of open data and data sharing initiatives by public and private actors (2022) p. 50 [online]. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf).

See also Commission Staff Working Document, Evaluation of Directive 96/9/EC on the legal protection of databases, SWD(2018) 146 final, 2018, p. 36: “In the context of automated data collection by sensor-equipped, connected ‘Internet of Things’ objects, it becomes increasingly difficult to distinguish between data creation and obtaining of data when there is systematic categorisation of data already by the data-collecting object (e.g. industrial robots).”

⁵³⁹ Matthias Leistner, ‘Big Data and the EU Database Directive 96/9/EC: Current Law and Potential Reform’ in S Lohsse, R Schulze and D Staudenmayer (eds), Trading Data in the Digital Economy: Legal Concepts and Tools (Hart Publishing 2018) 28.

⁵⁴⁰ Marta Duque Lizarralde, “4iP Council - A Guideline to Artificial Intelligence, Machine Learning and Intellectual Property” (2020) p. 7 [online]. Available at: https://www.4ipcouncil.com/application/files/9016/0017/8691/A_Guideline_to_Artificial_Intelligence_Machine_Learning_and_Intellectual_Property.pdf.

⁵⁴¹ Database Directive, Article 7(1).

⁵⁴² Database Directive, Recital 46: “Whereas the existence of a right to prevent the unauthorized extraction and/or re-utilization of the whole or a substantial part of works, data or materials from a database should not give rise to the creation of a new right in the works, data or materials themselves.”

⁵⁴³ Matthias Leistner and Lucie Antoine, IPR and the use of open data and data sharing initiatives by public and private actors (2022) p. 57 [online]. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf).

Under this sub-chapter, it is also important to refer to the recent proposal for the Data Act, where the EU's intention to move away from a more property-based approach to data and its protection is evident, with the goal to "incentivize the opening-up of privately held databases" and to "preserve a competitive environment and avoid monopolistic situations in information markets".⁵⁴⁴ Article 35 of the proposed Data Act illustrates the Act's aim to avoid countering "the extension of proprietary claims to data" by excluding the application of the sui generis database right to "databases containing data obtained from or generated by the use of a product or a related service"⁵⁴⁵, which falls within the scope of the proposed act. This concerns primarily databases containing IoT data and, in general, machine generated data. The reason for the exclusion of such databases from the scope of the sui generis database right is given under the recitals of the act as the failure in fulfilling the substantial investment criteria, where the substantial investment, as discussed above, is expected to be made in either the obtaining, verification or the presentation of the data.⁵⁴⁶ Overall, depending on the type of database and the content included therein, which the data provider wishes to allow access to, the sui generis database protection may become applicable or not. Any rights stemming from the sui generis protection the data provider chooses to assert can be considered as a part of the agreement manager within TRUSTEE. Nevertheless, where sui generis database protection is available, the limitations to and legitimate uses of this right recognized under the relevant EU legislations (such as the Database Directive and the recent proposed Data Act) as well as national laws must be considered by the parties entering into an agreement.

TRADE SECRETS / KNOW-HOW

Due to the limited means of protection of datasets by the above-mentioned intellectual property rights, trade secret protection is quite often a resorted means "for both the protection and the trading of data."⁵⁴⁷ This is due to the encompassing nature of the term 'trade secret', under which potentially "all core elements of a data-driven economy, such as diverse forms of datasets (including meta-data, excluded however a single datum)"⁵⁴⁸ can fall, as well as "co-generated data and individual use data"⁵⁴⁹ since combining and aggregating large sets of data can increase their commercial value. The emphasis is again on "datasets" rather than "individual data" since, though this might differ from sector to sector, individual data might be deemed as commercially valuable "only if they are part of larger datasets."⁵⁵⁰ Despite the encompassing nature of the definition for trade secrets, the applicability of the main criteria for the protection might still prove to be difficult for data. The commercial value of data needs to derive from its secrecy. In a data economy, not only that such causality might be challenging to prove, but "establishing the steps

⁵⁴⁴ Margoni, Thomas, Ducuing, Charlotte, & Schirru, Luca. (2023). Data property, data governance and Common European Data Spaces. <https://doi.org/10.5281/zenodo.7867881>, p. 4: The authors interpret the EU's data strategy as a "rather explicit choice towards the abandonment of the traditional (intellectual) property arena as a regulatory device for data exchanges, in favor of a more granular data governance regime."

⁵⁴⁵ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on harmonised rules on fair access to and use of data (Data Act) (Text with EEA relevance) {SEC(2022) 81 final} - {SWD(2022) 34 final} - {SWD(2022) 35 final}, COM(2022) 68 final 2022/0047 (COD), available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0068>

⁵⁴⁶ Proposal for Data Act, Recital 84.

⁵⁴⁷ Matthias Leistner and Lucie Antoine, IPR and the use of open data and data sharing initiatives by public and private actors (2022) p. 58 [online]. Available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU\(2022\)732266_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/732266/IPOL_STU(2022)732266_EN.pdf).

⁵⁴⁸ Ibid.

⁵⁴⁹ Ibid.

⁵⁵⁰ Francesco Banterle, Data ownership in the data economy: a European dilemma, (August 1, 2018). EU Internet Law in the digital era, (edited volume based on the REDA 2017 conference), Springer, (2018), available at SSRN: <https://ssrn.com/abstract=3277330>, p. 7.

necessary to keep the data secret in a network environment might also be a difficult task.”⁵⁵¹ These are points that should be considered by data holders/providers while evaluating the trade secret status of their particular data/datasets. It should also be noted that even though “the Directive introduces a broad and effective right of control over data”, it still remains “neutral on the question of ownership of information.”⁵⁵²

The Trade Secret Directive regulates not only the unlawful acquisition, but also the unlawful use and disclosure of trade secrets/know-how, meaning to which the trade secret holder has not consented (i.e. acquisition of information through spying or theft). As within the framework of TRUSTEE, the data holders/Data Providers share their data with Consumers, the provisions of the Directive concerning unauthorized access and unlawful acquisition will not apply, as protection given to trade secrets is conditional on the information still being kept confidential. Any contractual limitations to the further use or disclosure of the trade secret shared as data/datasets through TRUSTEE, however, shall continue to apply and any breach of these will be considered as unlawful use or disclosure of the trade secret if the use or disclosure of it is in “breach of a confidentiality agreement or any other duty not to disclose the trade secret” or is in breach of a contractual or any other duty to limit the use of the trade secret.”⁵⁵³ In these types of situations, entitlement to damages and other legal remedies may apply.

However, where the acquisition, use or disclosure of a trade secret is required or allowed by Union or national law, such acquisition, use or disclosure shall deem to be lawful.⁵⁵⁴ Additionally, where the alleged acquisition, use or disclosure of the trade secret was carried out for the purpose of protecting a legitimate interest recognised by Union or national law, this shall provide as an exception to the trade secret protection.⁵⁵⁵ These two points should be considered when determining the access and usage rights to data which are classified as a potential trade secret by data holders.

Additionally, if the trade secret is made publicly available, protection can no longer be claimed as the information loses its value to the company holding it as a trade secret.⁵⁵⁶ Academics point out the necessity to conduct further research on “whether the data that are broadly shared through commercial transactions based on non-disclosure agreements should retain their trade secret status.”⁵⁵⁷

⁵⁵¹ Ivan Stepanov (2020) Introducing a property right over data in the EU: the data producer’s right – an evaluation, *International Review of Law, Computers & Technology*, 34:1, 65-86, DOI: [10.1080/13600869.2019.1631621](https://doi.org/10.1080/13600869.2019.1631621), p. 72.

⁵⁵² Legal study on ownership and access to data, Final report – Study, European Commission 2016, p. 10 [online]. Available at: <https://op.europa.eu/s/n2Qc>. See also p. 11: “although the Directive itself has no provisions as to licensing arrangements, businesses will be able to treat data as another form of asset, with clear potential for the exploitation of previously private datasets in similar ways or through licensing or franchising.”

⁵⁵³ Trade Secret Directive, Article 4(3).

⁵⁵⁴ Trade Secret Directive, Article 3(2).

⁵⁵⁵ Trade Secret Directive, Article 5(d).

⁵⁵⁶ Tanya Aplin, ‘Trading Data in the Digital Economy: Trade Secrets Perspective’ in S Lohsse, R Schulze and D Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Hart Publishing 2018) 66-67: “Whether data is secret will de-pend on whether it has been treated as an open or closed source and, in part, on its type. For those datasets that are part of the open data movement clearly these cannot be treated as secret. Other datasets, however, may not be openly accessible, in which case the question will be whether the data is not ‘generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question.’”

⁵⁵⁷ Josef Drexler, Reto M. Hilty et al., “Artificial Intelligence and Intellectual Property Law Position Statement of the Max Planck Institute for Innovation and Competition of 9 April 2021 on the Current Debate” p 4. [online]. Available at: https://www.ip.mpg.de/fileadmin/ipmpg/content/stellungnahmen/MPI_PositionPaper_SSRN_21-10.pdf: “By ‘broadly shared data’ we mean data distributed, for instance, on a non-exclusive basis but still under contractual restrictions concerning their use and further disclosure.” See also ⁵⁵⁷ Legal study on ownership and access to data, Final report – Study, European Commission 2016 [online]. Available at: <https://op.europa.eu/s/n2Qc>, p. 12: “The

As a final note, it is important to point out that the proposed Data Act also provides for some interaction with the trade secret framework. Article 4(3) of the proposed Data Act states that trade secrets shall be preserved and be disclosed only when all specific necessary measures envisaged under the Trade Secret Directive are taken in advance to preserve the confidentiality of the trade secrets. In situations where the data holder is not also the trade secret holder, they are tasked with the identification of the data, which are protected as trade secrets, and to agree with the users any technical and organizational measures to preserve the confidentiality of the shared data and on liability provisions. The technical and organizational measures are exemplified by the proposed act as “model contractual terms, confidentiality agreements, strict access protocols, technical standards and the application of codes of conduct.”⁵⁵⁸

A similar provision concerns the disclosure of trade secrets to third parties, yet comes with additional requirements. Article 5(8) of the proposed act states that trade secrets shall only be disclosed to third parties “to the extent that they are strictly necessary to fulfil the purpose of the request agreed between the user⁵⁵⁹ and the third party and all specific necessary measures agreed between the data holder and the third party are taken prior to the disclosure by the third party to preserve the confidentiality of the trade secret.”⁵⁶⁰

In consideration of these proposed provisions and the trade secret framework as a whole, where the datasets, to which access will be allowed to, can be qualified as trade secrets, the data provider needs to evaluate the implementation of the above mentioned measures, where appropriate. Thus, TRUSTEE should allow for the facilitation and the negotiation of these as well.

IPRS AND AI MODELS

Within TRUSTEE, “artificial intelligence models as a service” will be developed and offered to users (model providers and consumers), where model providers can share their AI models, and consumers can create their own ML workflows to test their own algorithms or to make use of algorithms / AI models provided by TRUSTEE.

Focusing primarily on AI/ML models for the purposes of TRUSTEE, it is important to first set within the project the definition and the building blocks of an AI model straight, in order to understand whether it can either as a whole or partially be subject to protection by any intellectual property rights or not. However, this is not an easy task due to the lack of uniformity on how ML models are defined. They have been regarded as, for instance, either “learning algorithms”, “AI computer programs” or “super-software”.⁵⁶¹ It is also argued that AI/ML models may also encompass “a file with data structure containing the weights arranged values” which reminds of “the concept

major limitation of the Trade Secrets Directive, from the perspective of a datadriven commercial entity, is precisely that it requires the data to be kept secret; protection will be lost if the data is made public at any stage. Any licence granted over the data implies the risk of loss of control if the licensee fails properly to protect the data.”

⁵⁵⁸ Proposal for Data Act, Article 4(3).

⁵⁵⁹ User here is understood as “a natural or legal person that owns a connected product or receives a related service or to whom the owner of a connected product has transferred, on the basis of a rental or leasing agreement, temporary rights to use a connected product or receive related services...” per Article 2(5) of the proposed Data Act.

⁵⁶⁰ Proposal for Data Act, Article 5(8).

⁵⁶¹ Begoña Glez Otero (2021). Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose? Max Planck Institute for Innovation and Competition Research Paper No. 21-02, p. 2 & 6: “Such a paradigm shift comes with changes in software design practices that affect the software system and its components, while increasing complexity to the whole process. Also, the fact that ML models are learners that generate an output by making inferences from data, makes their technical delineation complicated. Are ML models algorithms and should they be treated from a copyright perspective as algorithms? Are they a new type of algorithm that requires copyright to look at them differently? Are they computer programs?”

of a database” and may even require an analysis, though partially, of database right.⁵⁶² However, a definition and classification of such models are important, as these will determine whether any protection can be granted to them by IPRs, and the licenses a party can rely on in allowing usage rights to them.⁵⁶³

Additionally, along with the uncertainties with the definitions, the protection of algorithms⁵⁶⁴ and artificial intelligence and machine learning models within the European intellectual property law framework is also not entirely clear. This presents challenges in assessing how terms concerning their access and use should be drafted. This is due to the following considerations, which will be elaborated upon further below:

- (a) the exclusion of protection of algorithms and the uncertainty of protection of AI/ML models by the EU Software Directive;
- (b) trade secret protection appearing to be the preferred option for the protection of algorithms and AI models.

COPYRIGHT PROTECTION – GENERAL CONSIDERATIONS

In order for an algorithm or an AI model to be considered as a work protectable by copyright per Infosoc Directive, it needs to be within the protected subject matter and satisfy the originality criteria, being reflective of the author’s own intellectual creation. The latter requirement necessitates the involvement of human contribution, since in essence, copyright is referred to as an “author’s right”. Due to this, for instance, it is contested that unsupervised learning algorithms would not be protected by copyright as human involvement is deemed to be limited in their creation.⁵⁶⁵

Additionally, if the “realisation of a subject matter has been dictated by technical considerations, rules or other constraints which have left no room for creative freedom, that subject matter cannot be regarded as possessing the originality required for it to constitute a work”.⁵⁶⁶ This can be overcome and copyright protection can be gained if the author was still able to reflect his personality in the subject matter and express his free and creative choices, despite the technical consideration being dictated in the realization of the work.⁵⁶⁷ However, if the expression of the components of a subject matter are dictated by their technical function, “the different methods of implementing an idea are so limited that the idea and the expression become indissociable”⁵⁶⁸ and as copyright is not granted to “ideas, procedures, methods of operation or mathematical concepts”, no copyright protection would arise for such components.⁵⁶⁹ Stemming from this and in particular consideration of AI/ML models, any weights and parameters as part of the model, to the extent that they comprise ideas and principles and where their selection and expression are dictated by their technical function, may not be protected by copyright by themselves.

⁵⁶² Begoña Glez Otero (2021). Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose? Max Planck Institute for Innovation and Competition Research Paper No. 21-02, p. 23.

⁵⁶³ Begoña Glez Otero (2021). Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose? Max Planck Institute for Innovation and Competition Research Paper No. 21-02, p. 2: “In most open licenses, if the license is applied to something that is not protected by copyright (or related rights) the license is not triggered.”

⁵⁶⁴ Katarina Foss-Solbrekk, Three routes to protecting AI systems and their algorithms under IP law: The good, the bad and the ugly (2021) 16(3) Journal of Intellectual Property Law & Practice [online]. Available at: <https://academic.oup.com/jiplp/article/16/3/247/6143561>, p. 249: Though there seems to be a lack of consensus in literature on the definitions of algorithms, they can be described as “a method constituting a sequence of steps used to calculate different data variables to yield outputs.”

⁵⁶⁵ Foss-Solbrekk (n 358) 251.

⁵⁶⁶ *SI, Brompton Bicycle Ltd v Chedech/Get2Get* (C-833/18) [2020] ECLI: EU:C:2020:461, para. 24.

⁵⁶⁷ *Brompton Bicycle*, para 26.

⁵⁶⁸ *Brompton Bicycle*, para 27.

⁵⁶⁹ *Levola Hengelo BV v Smilde Foods BV* (C-310/17) [2018] ECLI:EU:C:2018:899, paras. 39-40.

COPYRIGHT PROTECTION FOR COMPUTER PROGRAMS/SOFTWARE

Analysing AI/ML models as some sort of a computer program/software and considering their protection accordingly, though still being heavily disputed, is a supported approach in literature. Though as previously mentioned, an AI/ML model might have various parts that might be worthy of IPR protection, when the software copyright protection can step in is briefly discussed below.

Firstly, if an AI/ML model is considered as an algorithm or consisting partly of an algorithm, this will not be subject to computer program/software protection as the EU Software Directive explicitly excludes the protection of algorithms from its scope.⁵⁷⁰ The case law of the Court of Justice of the EU also supports this, where the Court has affirmed that “the source code and the object code of a computer program are forms of expression”⁵⁷¹ and thus are protected by copyright as computer programs whereas “the functionality of a computer program nor the programming language and the format of data files used in a computer program in order to exploit certain of its functions”⁵⁷² do not constitute expressions of the computer program.

Although an AI/ML model may not be necessarily considered as a “regular algorithm”⁵⁷³ but a sort of an algorithm “based upon a nonlinear mathematical function, capable of generating output based on the learned patterns in the training process”⁵⁷⁴, ML algorithms would still be excluded from copyright protection, “to the extent that they comprise ideas and principles.”⁵⁷⁵ Nevertheless, though an “algorithm itself constitutes a computer functionality” and thus is not protected under the Software Directive, “the source code of an algorithm attracts copyright protection”.⁵⁷⁶ If the software code meets the originality requirements and is protected by copyright, the “training algorithms encoded in a programming language and embedded in software”⁵⁷⁷ may also be protected. Yet, this is heavily discussed in literature and there is no clear-cut answer.

Another difficulty in fitting the ML models into the framework of copyright protection of computer programs is the existence, in the legal conceptualization of computer programs, of a “focus on the end product of the computer program development process (the executable code alone) at the expense of other pieces which exist at intermediary stages and which in some sense determine the final product.”⁵⁷⁸ This is due to the potential qualification of ML models as “part of intermediary stages of the ML system and written in code to enable their readability by

⁵⁷⁰ Database Directive, Recital 11.

⁵⁷¹ *Bezpečnostní softwarová asociace – Svaz softwarové ochrany v Ministerstvo kultury* (C-393/03) [2010] ECLI:EU:C:2010:816, para 34.

⁵⁷² *SAS Institute Inc. v World Programming Ltd.*, (C-406/10) [2012] ECLI:EU:C:2012:259, para 39.

⁵⁷³ Begoña Glez Otero (2021). Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose? Max Planck Institute for Innovation and Competition Research Paper No. 21-02, p. 15.

⁵⁷⁴ *Ibid.*, p. 17.

⁵⁷⁵ *Ibid.*, p. 25.

⁵⁷⁶ Foss-Solbrekk, 250.

⁵⁷⁷ Carlos Muñoz Ferrandis and Marta Duque Lizarralde ‘Open sourcing AI: intellectual property at the service of platform leadership’ 13 (2022) JIPITEC 231.

⁵⁷⁸ Begoña Glez Otero (2021). Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose? Max Planck Institute for Innovation and Competition Research Paper No. 21-02, p. 23.

computers, but not necessarily self-executable.”⁵⁷⁹ Despite this, AI models may still be characterized as computer programs/software where a ML model “can be expressed in a coded form”.⁵⁸⁰

Upon this qualification however, a satisfaction of the requirements of protection granted to computer programs needs to be found. Apart from the originality requirement, which itself is challenging to satisfy and will require a case-by-case analysis, the finding of the necessary human element (as copyright authorship is mainly reserved to humans) in AI/ML models and systems might be difficult to prove. ML systems are said to be not exactly fitting of a “human creative software programming, as they do not generate human readable source code and because the role of a human programmer can be minimized in particular development phases to merely providing a list of desired inputs and outputs.”⁵⁸¹ Regardless of this, the involvement of a human in different tasks associated to ML models, such as feeding data for the completion of a task or implementation of a “detailed creative plan, which will define the key elements of the work” might still suffice to determine adequate human involvement.⁵⁸² Further research concerning the protection of AI/ML models and the different parts creating the model by copyright, specifically copyright protection for software/computer programs needs to be conducted.

TRADE SECRET / KNOW-HOW

Trade Secret Directive does not differentiate between types of information subject to protection, but rather conditions the information to be kept unknown to others as the first requirement to be regarded as a trade secret and also to have commercial value due to being kept secret. Due to the confusing nature of the protection from the traditional IP rights’ side, plenty of companies choose to protect their algorithms, software in which the algorithm is expressed, and the models by way of trade secrets. This is done so by way of restricting access to these both internally and externally, implementing technological measures to keep such information secure, signing non-disclosure agreements and licensing agreements in a restrictive manner. The same applies to raw data and datasets too, as mentioned in the earlier chapters. Where these are not protected by intellectual property rights, “companies often restrict access to them through contractual restrictions or technical protection measures, creating de facto control.”⁵⁸³

CONTRACT LAW AND LICENSING CONSIDERATIONS

The European intellectual property law framework, as discussed in above sub-chapters, may be presenting more issues than offering direct solutions in terms of the protection of data and AI models to which access is envisaged to be mediated via TRUSTEE. This is also outlined in the study conducted by the European Commission on legal ownership and access particularly to data as the following:

- “the analysis of each right depends on its own rules, making the analysis complex and multi-layered;

⁵⁷⁹ Ibid, p. 23. See also p. 24: “However, even if the CJEU has given a purposive interpretation of the Directive so that in order to exploit certain functions of a computer program, neither the functionality of a computer program nor the programming language and the format of data files used, constitute a form of expression of that program, the question of protection of certain parts of non-executable code as independent works of copyright remains unsettled.”

⁵⁸⁰ Ibid, p. 25: „the scope of protection would be limited to that concrete formulation or specific arrangement of the algorithms contained in the ML model, its inner structure.”

⁵⁸¹ Begoña Glez Otero (2021). Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose? Max Planck Institute for Innovation and Competition Research Paper No. 21-02, p. 19.

⁵⁸² Ibid. See also p. 20: Despite the finding of a human involvement, “the paradigm shift in software design practices, might have brought about challenges in regard to claiming authorship of ML models, depending on the architectures and training techniques employed.”

⁵⁸³ Carlos Muñoz Ferrandis and Marta Duque Lizarralde ‘Open sourcing AI: intellectual property at the service of platform leadership’ 13 (2022) JIPITEC 231.

- within any single country, the scope and extent of each right are uncertain⁵⁸⁴; and
- in the international context, legal outcomes may differ between countries. For example for the same fact pattern copyright or trademarks may be enforceable in one country but not another.”⁵⁸⁵

Therefore, for the purposes of TRUSTEE, according to the initial analysis, rights in relation to data (e.g. access rights, usage rights) and AI/ML models, regardless of whether they are subject to any intellectual property rights, can initially be determined by Data/Model Providers, TRUSTEE developers and Consumers on a contractual basis. Contractual arrangements can provide a better solution to navigate this complex legal landscape. Relying on contracts and licenses can provide the required flexibility in jointly determining the conditions for the use of accessed data and models, and deciding on limitations, especially for data and models that are not subject to intellectual property protection but are nevertheless under the control of certain parties. This is also fitting to the reality of the data economy, where certain parties are exercising factual control, also called as factual exclusivity⁵⁸⁶, over datasets and AI/ML models by way of either contracts or physical or technological measures to restrict access.⁵⁸⁷

A contractual arrangement should include considerations stemming from intellectual property framework as described above, the confidential nature of certain types of data and models (hence trade secret protection) as well as granting the interacting parties the ability to reflect any expected uses of shared data and models (permitted uses and modifications such as enriching the data with other datasets) and the assignment of IP rights in the output of the model/any created data/datasets based on the used data and models. This analysis should feed into the development of the “agreement manager” as one of the tools to be developed within the Security and Trust Manager of TRUSTEE, through which data and AI model agreements are planned to be executed. TRUSTEE could focus on the development of standards and model terms for data and models or incorporate terms and standards from existing open licenses⁵⁸⁸ to encourage open and FAIR sharing of knowledge. Similar efforts are mainly channeled to the creation of standards and model contract clauses for data sharing. TRUSTEE could also engage in an exercise akin to this where standardized or model agreements or terms could be explored particularly for the sharing of and providing access to AI/ML models for the efficient and trustworthy sharing of the models.

TRUSTEE project, in the next steps of developing especially its terms of use and the agreement manager layer under the Security and Trust Manager component (under Task 4.6 and deliverables of WP4), should consider the

⁵⁸⁴ Directives only allow a certain level of harmonization and certain provisions of the directives might be allowing member states to enact higher protection measures. Due to this, the extent of certain rights might differ from member state to member state.

⁵⁸⁵ Legal study on ownership and access to data, Final report – Study, European Commission 2016, p. 12 [online]. Available at: <https://op.europa.eu/s/n2Qc>.

⁵⁸⁶ Ivan Stepanov, Introducing a property right over data in the EU: the data producer’s right – an evaluation, *International Review of Law, Computers & Technology* (2020) 34:1, 65-86, DOI: 10.1080/13600869.2019.1631621, p. 73: “...the way data are recorded, kept and protected, it is often the case that the data holder exercises factual exclusivity. It is from this position of factual exclusivity that the data holder is able to negotiate contracts on data access with other parties.”

⁵⁸⁷ Francesco Banterle, Data ownership in the data economy: a European dilemma, (August 1, 2018). *EU Internet Law in the digital era*, (edited volume based on the REDA 2017 conference), Springer, (2018), available at SSRN: <https://ssrn.com/abstract=3277330>, p. 9.

⁵⁸⁸ Usually, to ensure the effectiveness of open source licenses, the existence of underlying intellectual property rights are required. See Begoña Glez Otero (2021). Machine Learning Models under the copyright microscope: is EU Copyright fit for purpose? Max Planck Institute for Innovation and Competition Research Paper No. 21-02, p. 11: “there is a vast supply of simple ML models under very permissive open source licenses such as Apache 2.0 and MIT. However, these licenses are based on the legal right of copyright. If a ML model is not protectable under copyright, the license is not triggered, and it would be in the public domain.” Though, as the above sub-chapters showed, the protection of AI/ML models and datasets IPRs is heavily debated, contractual terms and licenses can provide remedy to a certain extent.

following points for the drafting and development of a licensing/contractual framework governing the relationship between TRUSTEE system deployers and its users, i.e. model providers providing models onto TRUSTEE or data providers providing datasets, between whom TRUSTEE will intermediate access with Consumers:

- As it is not easy to determine who owns certain models and datasets due to their common availability under open licenses, TRUSTEE could potentially adopt a less restrictive and more open approach where the providers are encouraged to share their models via TRUSTEE with open licensing terms⁵⁸⁹. These can cover, for instance, uses of the model for certain purposes, such as for research and development of further models. Certain limitations can be brought, such as the obligation to attribute the name of the provider of the models, but this should be carefully considered as it is not easy to ascertain the rightholders for subject matter, whose protection by intellectual property rights is uncertain and heavily discussed to begin with. To avoid any issues with misrepresenting rightholders, TRUSTEE can choose to accept only those models that are under a public domain license, where anyone is free to use, distribute and modify the model without requiring a permission.
- Connected to the above scenario, it might also be that model and data providers would like to provide models and datasets to TRUSTEE, on which they wish to exercise more restrictive terms and control and thus allow for their limited subsequent use. TRUSTEE could, under the Security and Trust Manager layer, facilitate agreements to be arranged where the providers negotiate terms on:
 - o the allowed uses of models/datasets;
 - o the limitations/prohibited uses brought to the use of the models⁵⁹⁰ and datasets;
 - o the terms regulating the results obtained by running the models or by adapting the models by running certain datasets with them (e.g. what sort of license should apply to the derived/adapted models or results of the models, who else can use these results and derived/adapted models⁵⁹¹).

The question of what sort of license and contractual terms should apply to what sort of situations is not straightforward to answer. This might also differ according to the different sectors to which the datasets and models are related to, where providers might have different expectations as to the contractual arrangements and guarantees they would like to see while sharing their datasets and models. This should be considered in line with the use cases of TRUSTEE and certain example licensing terms could be drafted in the later stages of the project accordingly, to

⁵⁸⁹ Javier de la Cueva and Eva Méndez, “European Commission Report of Study – Open Science and Intellectual Property Rights: How can they better interact? State of the art and reflections” (2022) p. 72 [online]. Available at: https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/open-science-and-intellectual-property-rights_en: “This generalisation, that free software is against copyright, is a common misunderstanding. Free software is another way authors have to exercise their copyright and then manage their intellectual assets. It does not mean that a work is free of copyright; it means that those rights are exercised in a different way from the traditional ‘all rights reserved’ paradigm. What these theories miss is that free software advocates the use of copyright to force the openness of their creations, which is a legitimate way to manage IP, often based on ethical considerations.”

⁵⁹⁰ As a further practices, the Responsible AI Licenses (RAIL), which include model terms for limiting use of AI models will be evaluated, more within the scope of work package 3. For further details on these licenses, see <https://www.licenses.ai/>. See also Paul Keller and Nicolò Bonato, Growth of Responsible AI Licensing. Analysis of License Use for ML Models Published on *Open Future* (February 2023) [online]. Available at: <https://openfuture.pubpub.org/pub/growth-of-responsible-ai-licensing>: “These licenses combine elements of permissive open source software licenses with provisions that restrict certain uses of the licensed resources. These restrictions are largely based on ethical considerations and are intended to ensure that the licensed models are used responsibly – hence the name Responsible AI licenses.”

⁵⁹¹ For example, as a consideration and in line with software copyleft licenses, rights concerning re-use of the derived/modified models can follow the same license terms applying to the initial models. If the initial model is made available under open-source licenses, this can encourage to follow open licenses for derived models as well.

test their application in the use cases. An exercise of engaging with the individual pilot use case leaders to explore with them the types of licenses and agreements they are currently using or witnessing the use of to license datasets and AI models can also be conducted within the project at later stages.

Next to this, as already mentioned above, it is important to prevent any false assertion of rights and excessive limitations to the use of the datasets and AI/ML models to which TRUSTEE facilitates access. There is a difficulty in accurately ascertaining whether the datasets, AI/ML models and workflows are protected by any intellectual property rights and who the rightholders are. Terms of Use of TRUSTEE platform could provide a solution to this by way of specific terms obliging users to provide accurate information for rights clearance and arranging liabilities accordingly. This is to be considered in the project and elaborated on in detail, if appropriate, in the upcoming iterations.

Finally, it is also important for TRUSTEE to implement, next to the above-discussed contractual means, the necessary technical infrastructure to be able to implement the terms on usage but, more importantly, on limitations brought to the use of datasets and models in order to ensure ongoing oversight of the data and model providers on how their datasets and models are used via and within TRUSTEE in order to establish trust in the process of sharing and granting access to datasets and models for other parties.

HOMOMORPHIC-CAPABLE SELF-SOVEREIGN FRAMEWORK

HOMOMORPHIC ENCRYPTION FOR DATA PRIVACY AND USABILITY

CONCEPTS

Homomorphic Cryptography (HE) is a cryptographic system whose main characteristic is that it allows arithmetic operations to be applied on or between ciphertexts (CT). In other words, it allows results to be obtained from a set of encrypted elements, without having access to the source data in clear, as it is shown in Figure 7. To achieve this, HE relies on the use of homomorphisms, which are functions that preserve the operations that can be applied over some objects, if they maintain the same algebraic structure⁵⁹². This technology is not new but had computational limitations that did not allow its application⁵⁹³ until, in 2009, Craig Gentry proposed a new encryption mechanism in his doctoral thesis and called it Fully Homomorphic Encryption (FHE)^{594,595}, from there, several HE schemes have emerged.

⁵⁹² Rossman, B. (2008). Homomorphism preservation theorems. *Journal of the ACM (JACM)*, 55(3), 1-53.

⁵⁹³ Rivest, R. L., Adleman, L., & Dertouzos, M. L. (1978). On data banks and privacy homomorphisms. *Foundations of secure computation*, 4(11), 169-180.

⁵⁹⁴ Gentry, C. (2009). A fully homomorphic encryption scheme. Stanford university.

⁵⁹⁵ Gentry, C. (2009, May). Fully homomorphic encryption using ideal lattices. In *Proceedings of the forty-first annual ACM symposium on Theory of computing* (pp. 169-178).

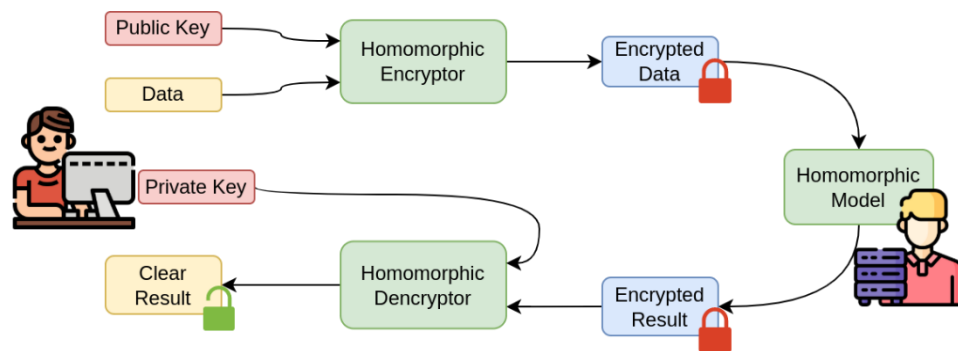


Figure 7: Schematic of the application of HE

TYPES OF HE SCHEMAS

When proposing a solution there are two things that must be considered when assessing which HE framework to use depending on the use case. It is important to properly choose the scheme that is going to be used to encrypt, there are several depending on the type of data and the operations needed to be computed in a homomorphic plane. Once the scheme that meets the requirements is clear, the next thing would be to assess which of the libraries best suits the needs of the use case.

The first categorization of schemas that can be done is based mainly on the number of iterations that are performed on the encrypted elements, or ciphertext (CT). These are:

Partially Homomorphic Encryption (PHE)

The cryptographic systems in this subset are those in which only one type of homomorphic operation (usually addition or multiplication) can be performed an unlimited number of times. These cryptography mechanisms are generally the least novel and least expensive of those to be described, and among them are RSA (Multiplicative), El-Gamal (Multiplicative) or Paillier (Additive)⁵⁹⁶.

Somewhat Homomorphic Encryption (SWHE)

In this case, cryptographic systems allow different types of homomorphic operations but only a limited number of times, since each iteration increases the noise of the ciphertexts, which makes the results end up being indecipherable after several operations. One of these mechanisms is called Boneh-Goh-Nissim (BGN)⁵⁹⁷.

Fully Homomorphic Encryption (FHE)

FHE encryption schemes are the most powerful and, in turn, the most limited today. They are the most flexible in terms of number of operations and iterations, but the most computationally expensive. These schemes allow an unlimited number of arithmetic operations, however, the increase in noise in each iteration is still present, which makes the result increasingly imprecise. To avoid this, noise reduction mechanisms, such as bootstrapping, are performed to improve the accuracy of results in models with multiple iterations⁵⁹⁵.

Levelled Fully Homomorphic Encryption (LHE)

⁵⁹⁶ L. Morris, «Analysis of partially and fully homomorphic encryption,» Rochester Institute of Technology, pp. 1-5, 2013.

⁵⁹⁷ D. G. E. J. & N. K. Boneh, «Evaluating 2-DNF formulas on ciphertexts,» de Theory of cryptography conference, Heidelberg., 2005

Finally, LHE schemes are cryptographic systems that inherit from FHE in which it is necessary to enter an L parameter, prior to encryption, which helps optimize the encryption and decryption processes to minimize their computational cost. This parameter determines the number of multiplications that can be performed on a ciphertext and has a direct impact on the size of the homomorphic keys, and therefore, on the sizes of the ciphertexts.

The main schemes that apply today are of the LHE type (BGV, BFV, CKKS, etc.), but in most documents and articles they are referred to as FHE, and this is also done in this document.

TYPES OF COMPUTATIONAL MODELS

The second categorization that can be made of homomorphic ciphers is the type of arithmetic they allow to be carried out on the encrypted data. Typically, the computational model required is a decisive criterion when dealing with a technical solution of this type.

Modular arithmetic

Modular arithmetic is an arithmetic system that uses congruences, or equivalences, to define properties between integers. In this type of operation, a value A can be congruent with B when, dividing it by the same value n, the same residue is obtained. This operation is called module n (mod), and an example of this is the calculation of the letter of the Spanish Identity Card, where module 23 of an 8-digit number is calculated, and the residue obtained is represented by one of the 23 letters obtained from a given equivalence table. The main drawback is that this system can only be operated with integer values.

Floating-point arithmetic

Floating point is a notation used by digital systems, composed of the elements exponent, sign and mantis, which is used to represent real numbers extremely efficiently, including numbers with decimals, although in some implementations of ciphers it is also used to represent numbers with two components, one real and one imaginary.

Boolean algebra

Boolean algebra encompasses all the logical operations on which the world of computing and digital electronics is based. Boolean circuits are used to represent it, composed of logic gates that make decisions based on binary or Boolean input data.

FHE ENCRYPTION SCHEMES

The encryption schemes mainly affect security, computational cost and their limitations in terms of implementation.

BGV / BFV

Brakerski, Gentry and Vaikuntanathan (BGV)⁵⁹⁸ is the most powerful encryption system that makes use of modular arithmetic to generate homomorphic ciphers. It is a very relevant FHE because it allowed to reduce the size of the public keys without using Gentry's bootstrapping noise reduction technique. When working with a modular arithmetic scheme, only integer values can be encrypted. This scheme only can compute arithmetic operations over integer numbers.

⁵⁹⁸ Z. Brakerski, C. Gentry, V. Vaikuntanathan, “(Leveled) fully homomorphic encryption without bootstrapping,” ACM Transactions on Computation Theory (TOCT), vol. 6, n° 3, pp. 1-36, 2014

Brakerski, Fan and Vercauteren (BFV)^{599,600} is a variant of BGV.

CKKS

Cheon-Kim-Kim-Song⁶⁰¹ is the most common floating-point arithmetic FHE encryption system, it works with an approximation arithmetic structure, which means that decryption is not an exact process, but an approximate value is deciphered. This technique can be used for arithmetic operations on decimals, and in the case of some frameworks, also with imaginary values.

TFHE

Fast Fully Homomorphic Encryption over the Torus⁶⁰² is an encryption system for evaluating homomorphic functions based on FHE, which uses a particular "gate-by-gate" bootstrapping technique to evaluate Boolean circuits composed of logic gates faster and more efficiently.

MULTIKEY FHE

A possible option in HE is to generate several public keys that can be used to generate interoperable ciphertexts. This mechanism is still under development, and it is not mature enough to be implemented in a project at this stage, but it is possible to use it to create proofs of concept.

In 2012, the University of Texas, together with Harvard University, created a multi-user HE protocol⁶⁰³, transforming a master key into multiple user keys. But they then use these keys to retrieve the data from a data-centric application, not to encrypt the input data themselves, in fact, their solution is based on a symmetric key that they have created themselves. Subsequently, there have been several studies around a multiple key protocol in homomorphic encryption (MKFHE)^{604,605,606,607}.

The closest there is to a tool that can be used to generate several homomorphic keys to operate on data from different sources is the Threshold FHE technique⁶⁰⁸, which is implemented in some libraries such as OpenFHE.

⁵⁹⁹ Z. Brakerski, "Fully homomorphic encryption without modulus switching from classical gapsvp," CRYPTO 2012, vol. 7417, pp. 868-886, 2012.

⁶⁰⁰ J. Fan, F. Vercauteren, "Somewhat practical fully homomorphic encryption," Cryptology ePrint Archive, Report 2012/144, 2012

⁶⁰¹ H. Cheon, A. Kim, M. Kim, Y. Song, "Homomorphic encryption for arithmetic of approximate numbers," de International Conference on the Theory and Application of Cryptology and Information Security, 2017

⁶⁰² I. Chillotti, N. Gama, M. Georgieva, M. Izabachene, "TFHE: Fast fully homomorphic encryption library," 2019

⁶⁰³ Xiao, L., Bastani, O., & Yen, I. L. (2012). An efficient homomorphic encryption protocol for multi-user systems. Cryptology ePrint Archive.

⁶⁰⁴ López-Alt, A., Tromer, E., & Vaikuntanathan, V. (2012, May). On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption. In Proceedings of the forty-fourth annual ACM symposium on Theory of computing (pp. 1219-1234).

⁶⁰⁵ Chen, L., Zhang, Z., & Wang, X. (2017, November). Batched multi-hop multi-key FHE from ring-LWE with compact ciphertext extension. In Theory of Cryptography Conference (pp. 597-627). Springer.

⁶⁰⁶ Chen, H., Dai, W., Kim, M., & Song, Y. (2019, November). Efficient multi-key homomorphic encryption with packed ciphertexts with application to oblivious neural network inference. In Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (pp. 395-412).

⁶⁰⁷ Ma, J., Naas, S. A., Sigg, S., & Lyu, X. (2022). Privacy-preserving federated learning based on multi-key homomorphic encryption. International Journal of Intelligent Systems.

⁶⁰⁸ Asharov, G., Jain, A., López-Alt, A., Tromer, E., Vaikuntanathan, V., & Wichs, D. (2012, April). Multiparty computation with low communication, computation and interaction via threshold FHE. In Annual International

Threshold FHE

The OpenFHE library provides a tool that applies the Threshold FHE technique⁶⁰⁸ on its homomorphic schemes. From this functionality it is possible to create several key pairs from a cryptographic context, so that it is possible for several users to have different interoperable keys, as it is shown in Figure 8.

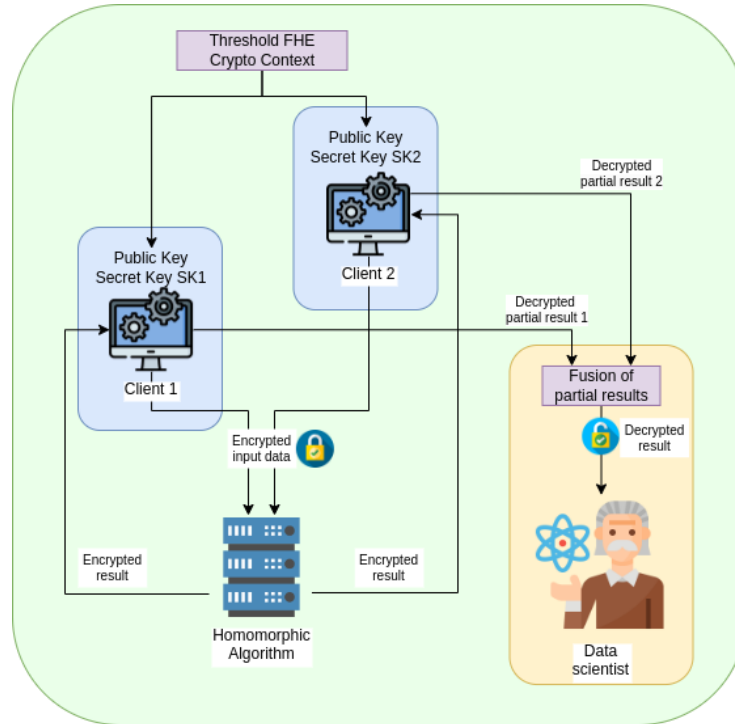


Figure 8: FHE Threshold operation

One of the advantages of this mechanism respect to the traditional FHE flow is that it is possible to use it for three or more users, in both BGV and CKKS. This technology can be very useful for the algorithm developer to verify the result after applying it, since it is not possible to decrypt the homomorphic result without the participation of all those involved in the process. But this also a drawback, because it requires a more complex crypto context (CC) configuration process, and it needs to be configured in a synchronous process, as can be seen in the Figure 9 for the case of two participants.

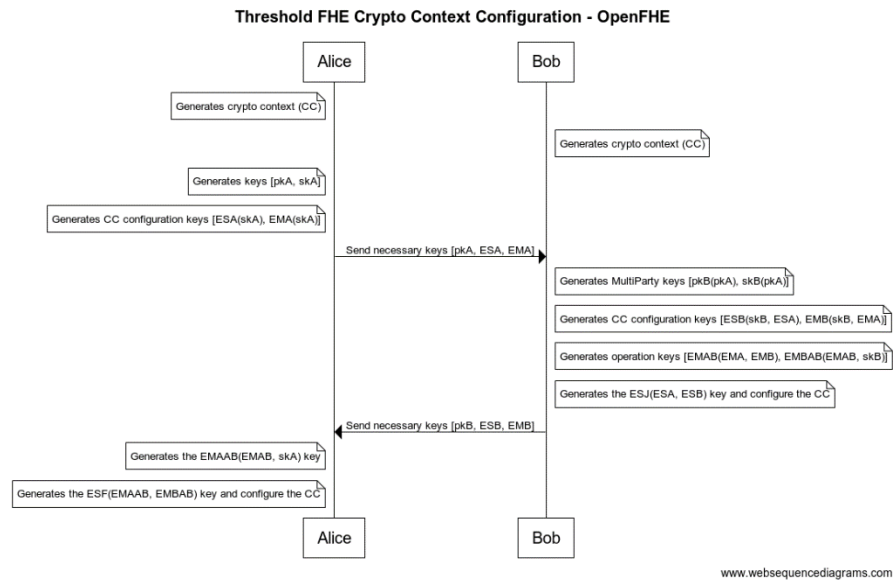


Figure 9: FHE Threshold configuration

The decryption process must be also synchronous, as shown in the Figure 10 for the case of two participants.

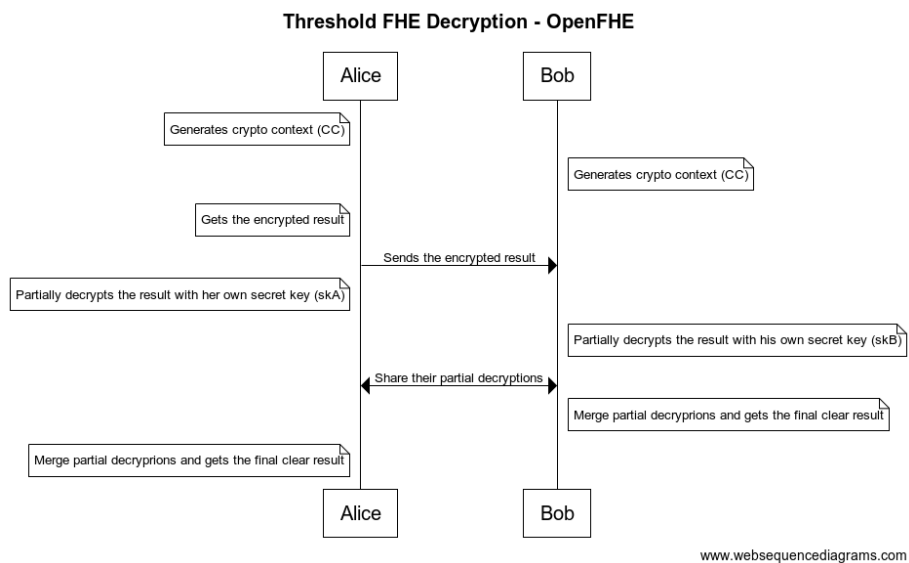


Figure 10: FHE Threshold decryption process

In addition, there is a major disadvantage when applying it in a context, which is that the input data cannot be reused unless the participants are not modified, since the user keys are bound to the cryptographic context used for encryption. Figure 11 shows an example of Threshold technology in OpenFHE for CKKS scheme.

```

=====RUNNING FOR CKKS=====
p = 50
n = 8192
log2 q = 210
Running key generation (used for source data)...
Round 1 (party A) started.
Round 1 of key generation completed.
Round 2 (party B) started.
Joint public key for (s_a + s_b) is generated...
Joint evaluation multiplication key for (s_a + s_b) is generated...
Joint evaluation multiplication key (s_a + s_b) is transformed into s_b*(s_a + s_b)...
Joint evaluation summation key for (s_a + s_b) is generated...
Round 2 of key generation completed.
Round 3 (party A) started.
Joint key (s_a + s_b) is transformed into s_a*(s_a + s_b)...
Computing the final evaluation multiplication key for (s_a + s_b)*(s_a + s_b)...
Round 3 of key generation completed.

Original Plaintext:
(1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1, ... ); Estimated precision: 50 bits
(1, 0, 0, 1, 1, ... ); Estimated precision: 50 bits
(2, 2, 3, 4, 5, 6, 7, 8, 9, 10, ... ); Estimated precision: 50 bits

Resulting Fused Plaintext:
(4, 4, 6, 9, 11, 12, 12, 12, 12, 12, 1, 1.04678e-07, ... ); Estimated precision: 23 bits

Resulting Fused Plaintext after Multiplication of plaintexts 1 and 3:
(2, 4, 9, 16, 25, 36, 35, 32, 27, 20, -6.66545e-08, 3.15568e-07, ... ); Estimated precision: 23 bits

Fused result after the Summation of ciphertext 3:
(56, 54, 52, 49, 45, 40, 34, 27, 19, 10, -9.15229e-08, -1.59127e-07, ... ); Estimated precision: 23 bits

```

Figure 11: Example of CKKS threshold technique execution using OpenFHE library

Some progress that are being made on this technology include attempts to make the process asynchronous⁶⁰⁹ using technologies like Multi-Party Computation, but it is still a configuration that must be done simultaneously, and the design of an alternative protocol called ThFHE⁶¹⁰ which allows an efficient Threshold FHE decryption over TFHE scheme to be applied on real-time systems.

TEXT SEARCH FHE

Text encryption using HE is a new research area that has seen little technological progress. Actually, the first appearance of this concept was in 1970 when Boneh proposed the Public-key Encryption Keyword Search (PEKS) scheme⁶¹¹, on which several improvements have been made over the years, but always maintaining the "sender, receiver, operator" system, which is a system that depends entirely on a trusted third party, which does not guarantee the integrity of the search results. It was not until 2018 when a new PEKS scheme was proposed, based on a multi-user technique⁶¹².

⁶⁰⁹ Urban, A., & Rambaud, M. (2022). Share & shrink: Ad-hoc threshold FHE with short ciphertexts and its application to almost-asynchronous MPC. Cryptology ePrint Archive.

⁶¹⁰ Chowdhury, S., Sinha, S., Singh, A., Mishra, S., Chaudhary, C., Patranabis, S., ... & Mukhopadhyay, D. (2022). Efficient FHE with Threshold Decryption and Application to Real-Time Systems. Cryptology ePrint Archive.

⁶¹¹ Kim, J. M., Jeong, H. Y., & Choi, H. G. (2012). Approach of Security Services on Smart-Phone using Homomorphism. In Applied Mechanics and Materials (Vol. 197, pp. 624-627). Trans Tech Publications Ltd.

⁶¹² Wu, D. N., Gan, Q. Q., & Wang, X. M. (2018). Verifiable public key encryption with keyword search based on homomorphic encryption in multi-user setting. IEEE Access, 6, 42445-42453.

Following this approach of homomorphic text ciphers, Ziroh Labs library is the only one that claims to be able to apply regular expression searches on encrypted data, but in this case, it is a private proposal of which not much data is known. There is also a tool developed in Java by Luke Thomas in 2018⁶¹³ where a demonstration of HE of numbers, texts and images is performed, but it does not apply a homomorphic search operation on this data, but it is limited to demonstrate that the properties of Shamir Secret Sharing are satisfied for all cases.

Regarding homomorphic searches, in 2021⁶¹⁴ a protocol was proposed for searching numerical values in which, given a set of encrypted numerical values, the ciphertext to be searched for is subtracted from all of them, and then all the resulting elements are multiplied between them. If the decrypted result is a zero, it means that the set of numbers contains the number that has been searched for, but it is not known which of all of them it is. This system would not be valid if one could iterate unboundedly over all the data in the set one by one.

Later on, in 2018 a protocol called Secure Search⁶¹⁵ was developed, also based on a cloud structure, which applies some multiple ring techniques together with a polynomial called SPiRiT (technique to compute the first positive index) to obtain this secure search.

Finally, also in 2018, a protocol named Zero Knowledge Set Membership (ZKSM) was developed, based on their previous protocol implementation Zero Knowledge Range Proofs (ZKRP), which can be used to verify that a given string element is contained within a set of further string elements, without revealing any information. This is quite similar to perform a string search over private data using Zero Knowledge Proof technology, with the condition that the text to be searched must be contained in the database where it is searched.

SUCCESS CASES

There are not many known success cases that use FHE so far. In 2018, the results of a European H2020 project named HEAT⁶¹⁶ were published, in which the HE tool HELib, developed by IBM, was applied on three different use cases: Satellites, Smart Grids and Automatic detection of organised crime, it is known that this tool has been further developed or applied but the specific cases are unknown. Also, in 2020, an article was published on the PRIORIS⁶¹⁷ application, which is used for early detection of suicidal intentions, and for this purpose, neural networks are homomorphically evaluated.

FHE LIBRARIES

Table 0.8 shows an overview of the main schemes of the last FHE generations.

Table 0.8 Summary of HE techniques.

Algorithm	TFHE	BGV/BFV	CKKS

⁶¹³ Thomas, L. (2018). Homomorphic Encryption Web-Based User Interface

⁶¹⁴ Kim, P., Jo, E., & Lee, Y. (2021). An Efficient Search Algorithm for Large Encrypted Data by Homomorphic Encryption. Electronics, 10(4), 484

⁶¹⁵ Akavia, A., Feldman, D., & Shaul, H. (2018). Secure search via multi-ring fully homomorphic encryption. Cryptology ePrint Archive

⁶¹⁶ THALES UK. (2018) HEAT (ICT-644209 - H2020): Homomorphic Encryption Applications and Technology: D4.3 Use Case Analysis and Validation.

⁶¹⁷ Natarajan, D., Dalskov, A., Kales, D., & Khanna, S. (2021). PRIORIS: Enabling Secure Detection of Suicidal Ideation from Speech Using Homomorphic Encryption. In Protecting Privacy through Homomorphic Encryption (pp. 133-146). Springer, Cham.

Computation	Boolean arithmetic	Modular arithmetic	Floating comma arithmetic
Data type	Booleans and Integers	Integers	Complex
Libraries	nuFHE, OpenFHE	Concrete, HELib, SEAL, OpenFHE	Lattigo, SEAL, Lattigo, OpenFHE, HELib

Table 0.9 and Table 0.10 describe the current status of some of the most relevant libraries for implementing FHE.

Table 0.9 Summary of the main FHE frameworks (17/01/23) [Part 1].

	LATTIGO	HElib	SEAL
Developer	Tune Insight	homenc (IBM)	Microsoft
Language	Go	C++17	C++17
Version	V4.1.0 (22/11/22)	V2.2.2 (06/12/22)	V4.1.1 (10/01/23)
Schemes	BFV/BGV, CKKS	BGV, CKKS	BFV/BGV, CKKS
Multi-party	DBFV/DBGV, DCKKS (Threshold)	-	-
License	Apache 2.0	Apache 2.0	MIT

Table 0.10 Summary of the main FHE frameworks (17/01/23) [Part 2].

	OpenFHE (PALISADE)	HEAAN	TFHE-rs (Concrete)
Developer	OpenFHEOrg	Cryptography LAB in Seoul National Uni.	Zama
Language	C++11	C++	Rust
Version	V1.0.3 (17/03/23)	V2.1 (05/09/18)	V0.2.1 (19/04/23)
Schemes	BFV/BGV, CKKS (only reals), FHEW/TFHE	HEAAN scheme	TFHE variant

Multi-party	BGV, BFV, (Threshold, Reencryption)	CKKS - Proxy	-
License	BSD 2-Clause	Creative Commons	BSD 3-Clause Clear

FHE LIMITATIONS

STORAGE

As mentioned before, one of the major limitations of this technology is the large weight of the encrypted data. It varies depending on the scheme used and the security parameters introduced on the cryptographic context (CC), but in the best case it can be of the order of 10^4 , stored in binary⁶¹⁸.

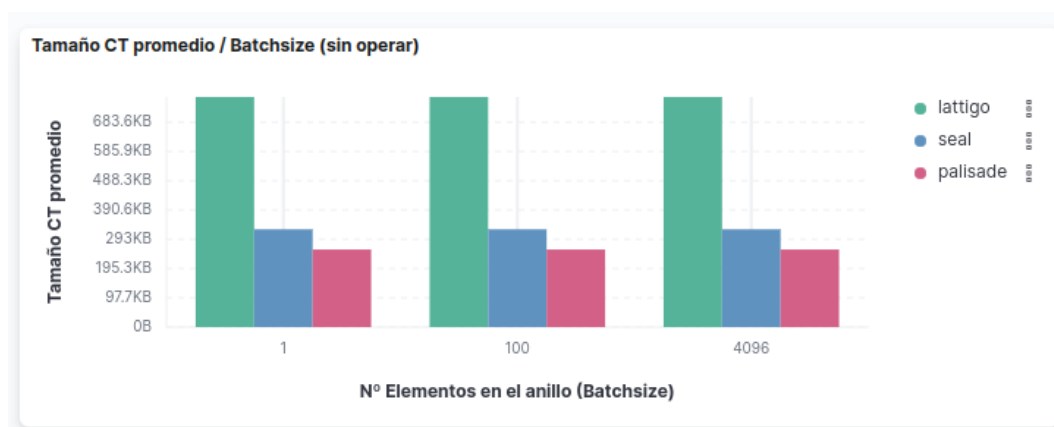


Figure 12: Average of CT size

For example, the size of a binary ciphertext (CT) of ring 4096, without having ever operated on it, is between 259KB and 7658KB as shown in Figure 12. We also know that the amount of data you enter the ring is independent of the size of the CT, since the leftovers are filled with zeros, but it does vary with the size of the ring that is configured.

It also increases the size of the CT depending on the operations that are done on it, but it depends on the library. For example, in OpenFHE, the context is configured to limit operations so that the process can be optimized, and that means that the size of the CT does not vary according to the operations.

OPERATIONS

⁶¹⁸ C. R. M. C. P. M. Á. P. a. E. J. Iñaki Seco Aguirre, "Comparative performance study of homomorphic cryptographic frameworks (in Spanish, "Estudio comparativo de rendimiento de frameworks de criptografía homomórfica")," de National Cybersecurity Research Conferences (in Spanish, "Jornadas Nacionales de Investigación en Ciberseguridad"), Bilbao, 2022

All libraries do not allow to perform any type of operation on any scheme. Table 0.11 gathers the main operations that can be performed. * *This table is an approach based on the documentation of each framework, which is not homogeneous for all cases. Not all operations have been tested so it may contain errors.*

Table 0.11 Operations implemented by the main FHE libraries

Basic Arithmetic	LATTIGO		Seal		Zama	OpenFHE		
	CKKS	BGV	CKKS	BGV	TFHE	CKKS	BGV	TFHE
CT Addition	x	x	x	x	x	x	x	
CT Subtraction	x	x	x	x	x	x	x	
CT Multiplication	x	x	x	x	x	x	x	
Negation	x	x	x	x	x	x	x	
Remainder					x			
PT add./subt.	x	x	x	x	x	x	x	
PT mult.	x	x	x	x	x	x	x	
Mult. by complex	x		x			x		
Slot rotation	x	?	x	x		x	x	
Square	x	x	x	x	x	x	x	
Square root						x		
Division by PT	x		x			x		
Bitwise					x			
Booleans					x			x
ShortInt comparisons					x			

Advanced Arithmetic	LATTIGO		Seal		Zama	OpenFHE		
	CKKS	BGV	CKKS	BGV	TFHE	CKKS	BGV	TFHE
Exponentiation	x		x	x				
Linear weighted sum						x	x	
Polynomial evaluation	x	x	x	x		x	x	
Inversion	x							
Linear Transformations	x					?		
Inner sum	x	x				x	?	
Inner product						x	?	
Replication (inner sum inverse?)	x							
Trace (diagonal?)	x							
Sum Rows						x	?	
Sum Columns						x	?	
Rotate columns/rows	x	x	x	x				
Number Theoretic Trans. (DFT)	x		x	?		?		
Galois automorphism	?	?	x	?		x	x	
Complex conj.			x					
Chebyshev approx.	?	?				x		

Sin					x	
Cos					x	
CT Division					x	
Logistic (sigmoid)					x	
Ciphertext management	LATTIGO		Seal		Zama	OpenFHE
	CKKS	BGV	CKKS	BGV	TFHE	CKKS BGV TFHE
Key-Switching	x	x	x	x	x	x x
Degree management	x	?	x	x	x	x ?
Scale management	x	?	x	?		x ?
Level management	x	?	x	x	?	x ?
Modulus reduction	x	x	?	?		x x
Object de/serialization	x	x	x	x	x	x x
CT Merge		x				x x
CT Compression			x	x		x x

In addition, in OpenFHE and Zama, which implement the TFHE scheme, boolean operations can be performed by encrypting bits. It should be noted that, in OpenFHE, the CC of TFHE, which allows operations with logic gates, is a symmetric key cryptographic system.

FUNCTIONAL

This section presents the functionality limitations that have been observed for HE using Lattigo.

In the case of BGV, BFV and CKKS, when an element or a set of elements is encrypted, it must be introduced into an array of elements of the size of the ring, which will later be transformed into PT and then into CT. And although, in a non-encrypted plane, the CT can be composed of several elements, once it is passed to a homomorphic encrypted plane, that is, to CT, all these elements that have been encrypted will become part of a single CT, and cannot be treated individually, in the homomorphic level you can only operate with the set of all of them, as explained in Figure 13.

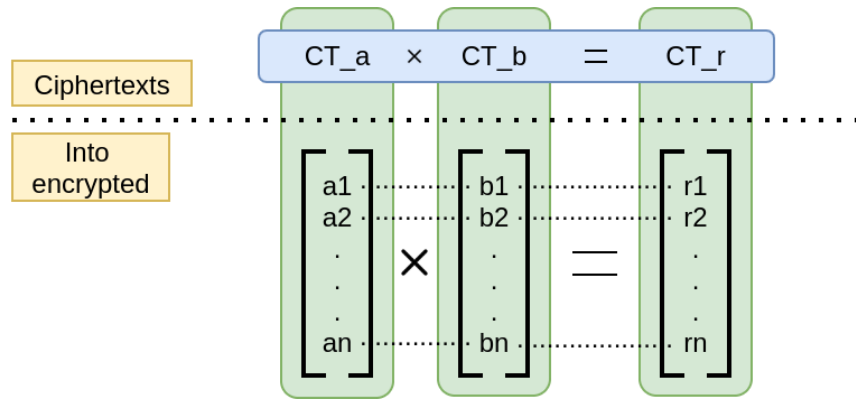


Figure 13: Schematic of an operation between two ciphertexts.

Generally, it cannot be performed operations between items that have been entered into the same CT. it can be performed operations only between the index n elements of each of the encrypted samples. There are some operations that allow you to operate between the elements inside a ciphertext, such as inner sum, but they are very rare.

There are limitations on the sizes of the array that are encrypted, this value corresponds to the size of the ring, and is a preconfigured value that varies depending on the security parameters used to encrypt. Unused gaps in the array will be filled with zeros, so that the size of the TCs does not change.

In this way, the nature of the input data determines, to a large extent, the rate of increase in size of the encrypted data. For example, successful proofs of concept in genomics in which arrays of a few thousand sections of the genome are processed by FHE, for a thousand individuals, with requirements of a few GB of memory in a few seconds. Sections of the genome are equivalent for everyone. However, conditions are more unfavourable for more samples of fewer elements (uncorrelated to each other but correlated with the equivalents of other samples).

These constraints refer only to the memory required for the management of the input data, which, a priori, seems to be the most determining factor. The verification of running memory consumption due to other factors (compute engine, intermediate variables, etc.) is still pending. Memory consumption seems a more limiting factor than storage capacity or download volume, although these conditions should not be ignored.

TIME

Figure 14, Figure 15 and Figure 16 and show the encryption times. The machine used was a Dell Latitude 5590 64-bit, Linux OS, 16GB RAM, Intel(R) Core (TM) i7-8650U CPU @ 1.90GHz.

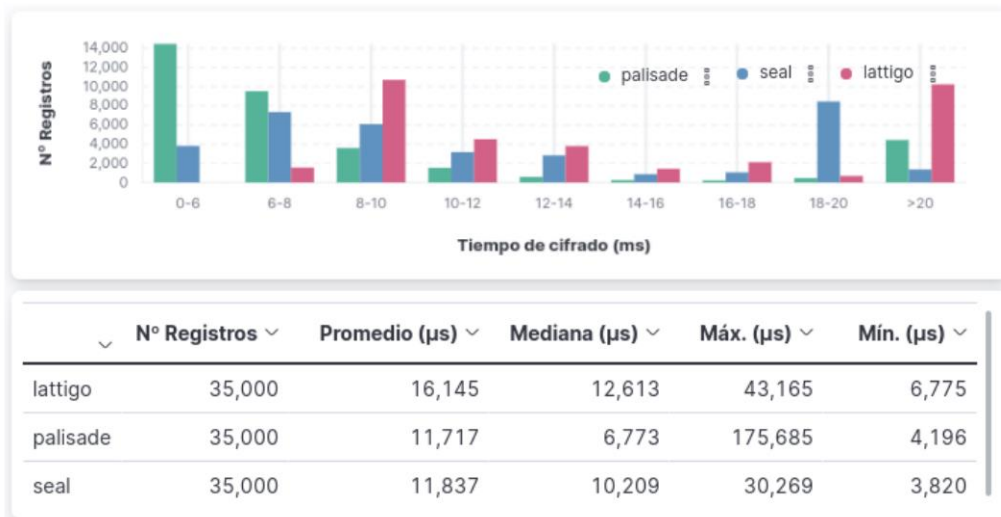


Figure 14: Encryption times for different frameworks in CKKS

	Nº Registros	Promedio (µs)	Mediana (µs)	Máx. (µs)	Mín. (µs)
lattigo	12,602	491	361	2,721	214
palisade	12,602	20,509	10,459	176,776	4,612
seal	12,602	491	330	2,608	94

Figure 15: Decryption times for different frameworks in CKKS

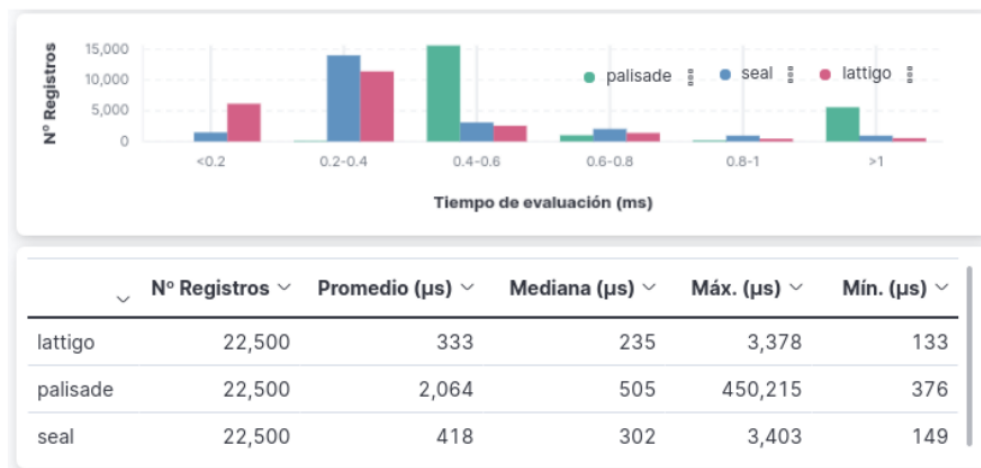


Figure 16: Evaluation times for different frameworks in CKKS

You can see from the figures that the slowest operation is decryption but considering that the number of iterations is probably much higher when encrypting, it should be noted that each encrypted element in CKKS can be around 12 ms, which can be an inconvenience for large amounts of data.

ACCURACY

There is also inaccuracy in the data once decrypted. Hundreds of tests can be done in this aspect, but preliminarily the test has been done to perform the sum operation between 0 and 10,000 times on the same CT, in intervals of 100 sums and deciphering after each interval, to assess the error committed according to the number of iterations on the same CT. The encryption scheme used is CKKS with a ring of 4096. The results are shown in Figure 17.



Figure 17: Increase of the error depending on the number of sums on a TC in CKKS

It can be observed that the error increases as you operate more times on the CT, and although the error is very small, it is not so small if we take into account that the encrypted values are decimal numbers between 0 and 1. The higher the encrypted value, the greater the error, which makes us conclude that the error made is proportional to the encrypted value, even so, we are talking about very small proportions.

SUMMARY

Homomorphic schemas influence the type of data you work with and how you operate with the data. BGV/BFV schemes are used to trade integer-only values and are not rendered inaccurate when a value is deciphered. The CKKS scheme is the most widespread and the one with the most operations implemented, it allows to operate with fractions and imaginary numbers, but the encryption and decryption operations introduce noise, which causes inaccuracies in the decrypted data, although cryptographic systems can be configured to minimize the error.

The TFHE scheme is used to operate with Booleans, but the Zama variant allows certain operations with integer values to be performed as well. They also allow you to execute operations with logic gates. This scheme is usually a symmetric cryptography system, but there is a chance to create a public key when use Zama as well.

The limitations are established mainly by the operations (Table 0.11) and the number of iterations which can be made over a CT.

In the case of BGV, BFV and CKKS, when an element or a set of elements is encrypted, it must be introduced into an array of elements of the size of the ring, which will later be transformed into plaintext and then into ciphertext, as shown in Figure 17. There are very few operations that allow operating between elements of the same ciphertext.

IDENTIFICATION, AUTHENTICATION, AND AUTHORIZATION BASED ON SSI FRAMEWORK

IDENTITY EVOLUTION

Identify Management (IdM) is a set of policies and technologies that ensures that authorized users have proper access to the technological resources. It requires an entity called “identity provider”, who provides identities to the different users. Often there is a “service provider”, who provides services to the different users. IdM has suffered

from different transformation along the years⁶¹⁹ starting with the so-called centralized model, then evolving to the federated model, and finally to the decentralized model. First of all, in the centralized model, the identity provider and the service provider were the same entity. That means that the actor who offers the service and offers the identity are the same. However, this model has some serious drawbacks, such as the need of remembering many user/password combinations, less security with re-used passwords, and identity presented in many different places, popularly known as “identity silos”. These identity silos became a common target for attackers looking for stealing users’ identities. After this, the federated model came up. In this identity model, the identity provider and the service provider became separated, which means that the actor who provides the service is not the same as the actor who provides the identity. It presents some advantages against the centralized model. For example, several service providers can trust in a single identity provider, thus reducing the number of places where identities are stored. Consequently, users do not need to remember as many user/passwords as in the centralized model, nor they need to re-use these passwords. Although this model presents clear advantages against the centralized model, it does not solve the core problem: identities are still stored in many different places, which also become a common target for attackers. The decentralized model comes to solve these drawbacks, proposing that users store their own identities, instead of being managed by an identity provider. In this context, Self-Sovereign Identity (SSI), a form of decentralized model, solves these drawbacks.

Focusing on existent IdM solutions, it is important to highlight that most of them assume that users and service providers trust the identity provider, which sometimes is wrong and can make the internal networks easily compromised. Furthermore, as introduced, identity providers are vulnerable because they constitute a Single Point of Failure (SPoF) and are common targets for attackers, mainly phishing attacks⁶²⁰. Some existent IdMs are discussed in⁶¹⁹ and even compared to each other, which are: PRIMELife⁶²¹, SWIFT⁶²², DAIDALOS⁶²³, Kantara⁶²⁴,

⁶¹⁹ Zhu, X., & Badr, Y. (2018). Identity management systems for the internet of things: a survey towards blockchain solutions. *Sensors*, 18(12), 4215

⁶²⁰ Alpar, G.; Hoepman, J.H.; Siljee, J. The Identity Crisis Security, Privacy and Usability Issues in Identity Management. arXiv 2011, arXiv:1101.0427.

⁶²¹ PRIMELife—Privacy and Identity Management in Europe for Life, Available online: <http://primelife.ercim.eu>. Accessed 02/01/2023

⁶²² SWIFT—SecureWidespread Identities for Federated Telecommunications. Available online: <https://www.swift.com>. Accessed 02/01/2023

⁶²³ DAIDALOS—Designing Advanced Network Interfaces for the Delivery and Administration of Location Independent, Optimised Personal Services. Available online: <https://www.tssg.org/projects/daidalos/>. Accessed 02/01/2023

⁶²⁴ Kantara. Available online: <http://kantarainitiative.org>. Accessed 02/01/2023

Liberty⁶²⁵, FIDIS⁶²⁶, SAML⁶²⁷, Higgins⁶²⁸, OpenID⁶²⁹, Shibboleth⁶³⁰, STORK⁶³¹, PICOS⁶³² and Cardspace⁶³³. However, these IdMs have lots of drawbacks, mainly centralization. However, these IdMs have lots of drawbacks, mainly centralization. Table 0.12, extracted from⁶¹⁹, gathers these different IdM systems and cross them with some characteristics:

Table 0.12. Comparison of IdM systems

	Scalability	Interoperability	Mobility	Security & Privacy	User-Centric
PRIMELife(PRIME)			X	X	X
SWIFT(DAIDALOS)		X	X	X	X
Kantara(Liberty)	X	X	X	X	
FIDIS		X	X	X	X
SAML		X			
Higgins		X		X	X
OpenID	X				X
Shibboleth				X	
STORK		X		X	X

⁶²⁵ Liberty—The Liberty Alliance Project. Available online: <http://www.projectliberty.org>. Accessed 02/01/2023

⁶²⁶ Matyas, V.; Fischer-Hübner, S.; Cvrcek, D.; Švenda, P. The Future of Identity in the Information Society; Springer: Berlin, Germany, 2009

⁶²⁷ Security Services Technical Committee of OASIS—Security Assertion Markup Language (SAML) v2.0. Available online: http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security. Accessed 02/01/2023.

⁶²⁸ Higgins—Open Source Identity Framework. Available online: <http://www.eclipse.org/higgins>. Accessed 02/01/2023

⁶²⁹ Recordon, D.; Reed, D. OpenID 2.0: A Platform for User-centric Identity Management. In Proceedings of the Second ACM Workshop on Digital Identity Management, Alexandria, VA, USA, 3 November 2006.

⁶³⁰ Cantor, S.; Scavo, T. Shibboleth Architecture. Protoc. Profiles 2005, 10, 16.

⁶³¹ STORK—Secure idenTity acrOss boRders linKed. Available online: <https://joinup.ec.europa.eu/collection/secure-identity-across-borders-linked-stork/about>. Accessed 02/01/2023

⁶³² PICOS—Privacy and Identity Management for Community Services. Available online: <https://cordis.europa.eu/project/id/215056/es>. Accessed 02/01/2023

⁶³³ Bertocci, V.; Serack, G.; Baker, C. Understanding Windows Cardspace: An Introduction to the Concepts and Challenges of Digital Identities; Addison Wesley: Boston, MA, USA, 2007.

PICOS	X	X	X	X	X
Cardspace		X	X	X	X

CONCEPT

SSI is part of a new identity model, which is the decentralized model. With SSI, users store their identities in their own wallet, thus avoiding centralization and SPoF. As credentials are stored on the user's side, it also enables selective disclosure by using Zero-Knowledge-Proofs (ZKPs). ZKPs are cryptographic functions that allow someone to prove that he has certain attributes without disclosing these attributes. For example, they would allow to prove that someone is over eighteen, without disclosing his age.

SSI requires a shared Verifiable Data Registry (VDR) between all parties. When the words "shared" and "registry" are put together, often the word "Blockchain" also comes up. However, blockchain is not mandatory as a technology for implementing the VDR, but it is highly desirable due to its unique properties. Blockchain guarantees that the information stored on it has not been modified, and also presents a decentralized infrastructure, avoiding SPoFs. As Blockchain is becoming an increasingly popular technology, researchers and companies from all over the world are trying to adapt the Blockchain for their use cases, presenting different and original approaches. Benefits of Blockchain to enhance the security of the system have been widely discussed along the recent years⁶³⁴ mainly in terms of privacy (if it is combined with ciphering), integrity, availability, traceability and non-repudiation.

SSI uses the concept of "verifiable credentials", issued to the user by a competent and trustable Certificate Authority (CA), which is a digital equivalent to our normal passports, issued by government authorities, and stored in a safe place, showing them as a proof when a verifier requires them. These verifiable credentials are called "verifiable" because they are signed by its issuer. For example, when a passenger gets in a plane, he proves his identity, issued by a trusted authority (claim issuer). Then he opens his wallet and shows his passport to the security personnel (claim verifier). Therefore, a credential is the evidence of a fact, capacity, state... associated with an identity. Blockchain is commonly used to implement the VDR. With Blockchain, each public key can own now its own address, which is now named as Decentralised Identifier (DID)⁶³⁵, another standard coming from the W3C. DIDs provide with a standard mechanism for creating decentralized identifiers for everyone, allowing them to keep them under their own control in their own wallet, similarly as happens with verifiable credentials. A DID is stored in a VDR with a DID Document, which describes the subject of the DID, its associated public key, public credentials, network endpoints, authentication protocols and digital sign. The following figure, extracted from the W3C ⁶³⁶, sums up the basic SSI flow:

⁶³⁴ Zheng, Zibin & Xie, Shaoan & Dai, Hong-Ning & Chen, Xiangping & Wang, Huaimin. An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. 10.1109/BigDataCongress. 2017.85.

⁶³⁵ W. España, «Decentralized Identifiers (DIDs) v1.0,» [Online]. Available: <http://www.w3.org/TR/did-core/>. Accessed 02/01/2023

⁶³⁶ World Wide Web Consortium. (2019). Verifiable Credentials Data Model 1.0: Expressing Verifiable Information on the Web. <https://www.w3.org/TR/vc-data-model/#core-data-model>.

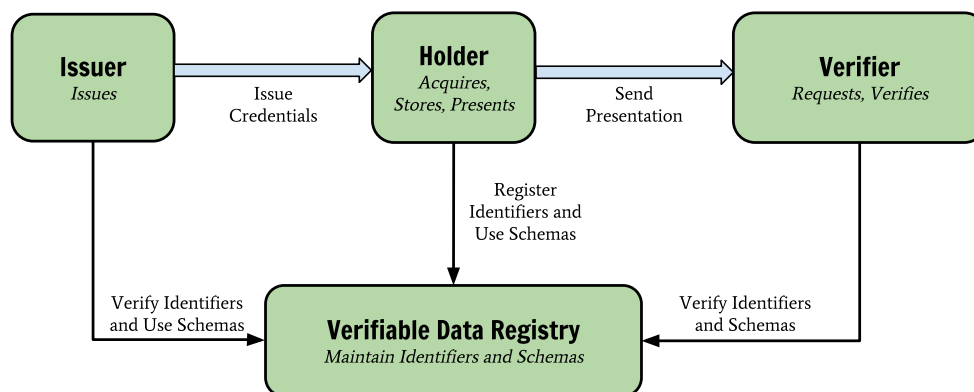


Figure 18: SSI flow

FRAMEWORKS

Current SSI technologies and frameworks include, but are not limited to:

- Hyperledger Indy (Sovrin)⁶³⁷: Hyperledger Indy is a permissioned Blockchain, which has been conceived directly for SSI, where anyone can read the content but only the interested parties can write on the ledger. Indy admits verifiable credentials, allowing to create ZKPs where a user can prove only the required attributes without revealing the full credential. The Hyperledger Indy adoption is being pushed by the Sovrin Foundation, which is an organization who works in establishing a trusted global network, which provides with legal basis and trust for SSI.
- Veramo⁶³⁸: Veramo provides a set of libraries and protocols for developing decentralized applications (dApps) focused on the identity over the Ethereum platform. It uses an interoperable identity layer, jointly with identity protocols, Smart Contracts and verifiable credentials. The library Ethr-DID⁶³⁹ provides with the capacity to any Ethereum address for gathering data in and out the chain, so any Ethereum address can become an identity using Veramo.
- Blockstack⁶⁴⁰: blockstack is a platform for the development of decentralized applications which provides a decentralized service of names and discovering where users can register readable resources and easy to find in the network. Consequently, it is not a tailor-made solution for SSI, but it implements a DID as a relationship between a username and his public key.
- SelfKey⁶⁴¹: SelfKey is an Open-Source identity management system built over Ethereum. It provides a user identities wallet, a decentralized marketplace with services and it uses tokens.

⁶³⁷ S. Foundation. «Sovrin: A Protocol and Token for Self-Sovereign Identity and Decentralized Trust» .[Online].Available: <https://sovrin.org/wp-content/uploads/Sovrin-Protocol-and-Token-White-Paper.pdf>. 2018. Accessed 02/01/2023

⁶³⁸ Performant and modular APIs for Verifiable Data and SSI. [Online]. Available: <https://veramo.io/>. Accessed 02/01/2023

⁶³⁹ Github. «Ethr-did» .[Online]. Available: <https://github.com/uport-project/ethr-did>. Accessed 02/01/2023

⁶⁴⁰ B. PBC. «Blockstack Technical Whitepaper v 2.0».[Online]. Available: <https://pdos.csail.mit.edu/6.824/papers/blockstack-2017.pdf>. 2017. Accessed 02/01/2023

⁶⁴¹ The SelfKey Foundation. «SelfKey».[Online]. Available: <https://selfkey.org/>. 2017. Accessed 02/01/2023

Other less known technologies for SSI are: Veres One⁶⁴², Jolocom⁶⁴³, Identity.com⁶⁴⁴, Uniquid⁶⁴⁵, Personal Identity⁶⁴⁶. The Decentralized Identity Foundation (DIF)⁶⁴⁷ gathers all of them.

There are also several frameworks enabling SSI in different regions and countries, being the most known:

- EBSI (European Union)⁶⁴⁸ : EBSI (European Blockchain Services Infrastructure) is a European platform for secure data sharing and enables communication between users within the European Union with their public administrations and companies. It is formed by a network of distributed nodes across Europe. Within this European network, one of the aspects is the identity management. The users of the framework are those who interact with the EBSI network and require its services and their identity must be correctly managed. To do so, EBSI provides with an identity framework, called ESSIF (European Self-Sovereign Identity Framework). This framework, based on the SSI paradigm, tries to establish some guidelines about how should be managed the identity of the European users using SSI schemes. More technically speaking, EBSI uses Hyperledger Fabric and Ethereum technologies for its implementation.
- PCTF (Pan-Canadian Trust Framework)⁶⁴⁹: From its side, the Canadian Government has also set up its own identity management framework, also following the standards of the SSI paradigm. Differently from the European framework, it does not specify the involved technologies used for the identity management process, but it also presents some guidelines about how this process should be made. Besides, it has ultimately the same objective than the European counterpart, which is to guarantee a safe and self-sovereign identity credentials exchange between Canadian citizens and their public administration and private companies. Also, within Canada, British Columbia⁶⁵⁰ is working on SSI for people.
- Distributed Identity Alliance (China)⁶⁵¹: In China, there has been a private initiative, which is the Distributed Identity Alliance, between some of the biggest companies in the country, to adopt standard mechanisms to manage the identity enabling a digital identity. This alliance intends to establish a cooperation between these companies to provide local business with a digital identity, also following the SSI paradigm.
- GAIA-X⁶⁵²: GAIA-X has presented its own whitepaper that focus on the importance of SSI in the GAIA-X ecosystem. The idea is to allow people, organizations or machines to manage their digital identities and

⁶⁴² «Veres One - A Globally Interoperable Blockchain for Identity» .[Online]. Available: <https://veres.one/>. Accessed 02/01/2023

⁶⁴³ «Jolocom - Own your Digital self» .[Online]. Available: <https://jolocom.io/>. 2019. Accessed 02/01/2023

⁶⁴⁴ «identity - Decentralized Identity Verification».[Online]. Available: <https://www.identity.com/>. Accessed 02/01/2023

⁶⁴⁵ «Uniquid . Connect and Control IoT Devices at Scale».[Online]. Available: <https://uniquid.com/>. Accessed 02/01/2023

⁶⁴⁶ « Personal Identity. Let users control and share their own verified data.» .[Online]. Available: <https://www.pingidentity.com/en/platform/capabilities/personal-identity.html>. Accessed 02/01/2023

⁶⁴⁷ «DIF».[Online]. Available: <https://identity.foundation/>. Accessed 02/01/2023

⁶⁴⁸ EBSI. Experience cross-borders services with EBSI.The first public sector blockchain services in Europe. [Online]. Available: <https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/Home>. Accessed 02/01/2023

⁶⁴⁹ Pan-Canadian Trust Framework (PCTF) – Digital Identity and Authentication Council of Canada (DIACC). [Online]. [Available](#) online. Accessed 02/01/2023

⁶⁵⁰ Verifiable Credentials for people. [Online]. Available: <https://digital.gov.bc.ca/digital-trust/projects-and-initiatives/credentials-for-people/>. Accessed 02/01/2023

⁶⁵¹ Distributed Identity Alliance led by FEITIAN has been launched. [Online] Available: <https://www.ftsaf.com/article/771.html>. Accessed

⁶⁵² Gaia-X secure and trustworthy ecosystems with Self Sovereign Identity. Developing a Decentralised, User-Centric, and Secure Cloud Ecosystem. [Online] Available: <https://www.gxfs.eu/download/3055/>. Accessed 02/01/2023

associated credentials (such as membership cards, certificates or self-descriptions) without having to rely on a conventional, central IdM system.

APPLICATIONS

Although SSI has been initially conceived for people, it can be extended to the Internet of Things (IoT). Some authors have proposed solutions that combine SSI and IoT. Some authors⁶⁵³ propose the use of a Smart Band for biometric identification following an SSI scheme, while other authors⁶⁵⁴ particularize the SSI paradigm for the IoV (Internet of Vehicles). In general, SSI can be applied to almost anything that has an identity that needs to be managed. For example, SSI could be used to identify datasets, as the project TRUSTEE proposes. This approach is not new as shown in some works⁶⁵⁵. In this work, SSI is presented as a tool to certify the authenticity of characteristics and provenance published datasets so that secondary users can inspect and gain confidence in the qualities of the data. The verifiable credentials are kept in the dataset's wallet and contain the hash of the dataset and a true/false value that represents that the dataset source has been certified or not. However, the paper lacks in providing details about the implementation and does not include a performance evaluation. In TRUSTEE, this model will be extended with more information and implementation details will be shared. TRUSTEE will also implement SSI for datasets within several industries, which are: automotive, education, energy, healthcare, space, etc.

However, SSI also has some worth-mentioning limitations. First of all, attributes are not confidential in the verifiable credentials nor in the verifiable presentation. An attacker could force access to the user's credential and read this information. Some technologies, such as Hyperledger Aries⁶⁵⁶, support storing ciphered verifiable credentials, but this information is revealed when it leaves the wallet. Furthermore, once a verifiable presentation has been created, attributes that have not been hidden through ZKPs can be read. TRUSTEE Project intends to apply ciphering using HE over these attributes to prevent this scenario. Also, ZKPs can only be applied to numeric attributes, so the project will try to extend ZKPs to non-numeric attributes, such as strings.

HE COMBINED WITH SSI

The Australian National University proposes the use of homomorphic cryptography to validate and obfuscate the data of a biometric signature⁶⁵⁷, which is used to access a Unique Self-Sovereign Identity (USI) network, where the credential data of the USI network is not encrypted. Related to this topic, a Revocable Identity-Based Fully Homomorphic Signature (RIBFHS) scheme was also developed⁶⁵⁸ and used to revoke privileges associated to an identity from untrusted computers, however, this technology is oriented to be used in a traditional user identity architecture, not based on blockchain.

⁶⁵³ Abishek, S. V., Priyan, G. D., More, S., & Suganthan, A. IoT based Smart Band for Biometric Authentication Using Blockchain Technology.

⁶⁵⁴ Theodouli, A., Moschou, K., Votis, K., Tzovaras, D., Lauinger, J., & Steinhorst, S. (2020, June). Towards a Blockchain-based Identity and Trust Management Framework for the IoV Ecosystem. In 2020 Global Internet of Things Summit (GloTS) (pp. 1-6). IEEE

⁶⁵⁵ Barclay, I., Radha, S., Preece, A., Taylor, I., & Nabrzyski, J. (2020). Certifying provenance of scientific datasets with self-sovereign identity and verifiable credentials. arXiv preprint arXiv:2004.02796.

⁶⁵⁶ N. George and et.al, "Hyperledger Aries Proposal," 2019. [Online]. Available:

<https://wiki.hyperledger.org/display/TSC/Hyperledger+Aries+Proposal>. [Accessed 2022].

⁶⁵⁷ Hamer, T., Taylor, K., Ng, K. S., & Tiu, A. (2019). Private Digital Identity on Blockchain. In BlockSW/CKG@ISWC

⁶⁵⁸ Xie, C., Weng, J., Lu, W., & Hou, L. (2020). General construction of revocable identity-based fully homomorphic signature. Science China Information Sciences, 63(3), 1-3.

IOTA is a distributed ledger technology designed to record and execute transactions between machines in the Internet of Things (IoT) ecosystem. The project's main innovation is Tangle, a system of nodes used for confirming transactions. However, IOTA is more than just a cryptocurrency; it's a technological framework with several unique features:

- **Tangle:** Unlike blockchain-based cryptocurrencies like Bitcoin and Ethereum, IOTA utilizes a directed acyclic graph (DAG) called the Tangle for its core protocol. Each transaction added to the Tangle is linked to two previous transactions, which creates a web of interconnected transactions. This structure allows for high scalability, no transaction fees, and a potentially high degree of decentralization.
- **No Mining:** Since all network participants in IOTA are equal, there are no miners to reward for confirming transactions. Instead, every user that wants to execute a transaction must approve two other transactions on the network. This process removes the need for transaction fees and ensures that the network can run even on devices with minimal resources, like sensors or other IoT devices.
- **Data Integrity:** IOTA is built to ensure data integrity and offers a secure way to verify the state of data at any given point in the network's history. This is particularly crucial for IoT devices, which often need to verify the accuracy and timeliness of data.
- **Quantum Resistance:** IOTA claims to be resistant to quantum computer attacks, which could potentially break the encryption methods used in other blockchains. This is an important feature as quantum computing continues to progress.

IOTA TECHNICAL DETAILS

THE IOTA 1.5 (ALSO KNOWN AS CHRYSALIS)

This was a major upgrade in the IOTA network that was implemented to improve the network's efficiency, reliability, and usability. It was seen as a transition before the full implementation of IOTA 2.0 (Coordicide). This comprehensive upgrade was critical for IOTA to achieve its goals of serving as a fundamental protocol for the Internet of Things (IoT), enabling feeless microtransactions, secure data transfer, and a host of other features necessary for IoT devices to communicate and transact with each other in a decentralized, autonomous way. The Chrysalis upgrade encompassed several key components:

1. **New Tip Selection Algorithm:** The tip selection algorithm, which decides which two transactions a new transaction will confirm, was overhauled to be more efficient and reliable.
2. **Binary Transactions:** Transactions and messages became binary rather than trinary, meaning they are more compatible with common hardware and network protocols, and more efficient to process
3. **Atomic Transactions:** The bundle concept from IOTA 1.0 was replaced with simpler atomic transactions, streamlining transaction processing and reducing room for user error.
4. **Reusable Addresses:** While IOTA 1.0 addresses could not be safely reused after funds were spent from them due to IOTA's use of Winternitz one-time signatures, Chrysalis introduced a new signature scheme that allowed addresses to be reusable, making the protocol easier to use correctly and safely
5. **UTXO model:** The Unspent Transaction Output (UTXO) model was adopted for tracking ownership of IOTA tokens. This model is more standard across the cryptocurrency space, with Bitcoin being a notable user of the model, and it has various advantages, including better handling of conflicts and double spends, and improved possibilities for future features such as smart contracts
6. **Performance improvements:** The IOTA Foundation claimed substantial improvements in transaction confirmation time, node synchronization speed, and overall network performance due to these changes.

THE IOTA 2.0 (ALSO KNOWN AS COORDICIDE)

This is a planned update aimed at removing the Coordinator, a special node run by the IOTA Foundation for network protection. This is to achieve a fully decentralized network. It represents a shift towards a fully decentralized network by removing the Coordinator, a special node run by the IOTA Foundation that was previously used to secure the network. The Coordinator was originally implemented as a protective measure against attacks, but it was also a point of centralization within the Tangle, as it was responsible for validating and confirming transactions. The objective of Coordicide is to decentralize this aspect of the network, ensuring that the Tangle can operate securely without relying on a central authority. Coordicide brings a set of new features and improvements to the IOTA network:

1. **Sharding:** Sharding is a scaling solution that splits the Tangle into multiple, smaller Tangles (shards), each capable of processing transactions independently. This allows the network to scale horizontally, improving its capacity as the number of users and transactions grows.
2. **Mana:** Mana is a new reputation system introduced in IOTA 2.0. It's a resource that nodes gain when they process transactions. Nodes with more mana have more influence in the network, encouraging good behavior and making it costly for bad actors to harm the network.
3. **Autopeering:** Autopeering is a feature where nodes automatically find and connect to other nodes in the network. This strengthens the decentralization and resilience of the network, as it minimizes the risk of nodes being isolated or the network being partitioned.
4. **Fast Probabilistic Consensus (FPC):** FPC is a new consensus mechanism designed to be lightweight and efficient, enabling the network to quickly agree on the state of transactions.
5. **Integration of Smart Contracts:** With the launch of IOTA 2.0, IOTA plans to introduce a new smart contract platform. This will enable developers to create and deploy decentralized applications (dApps) on the IOTA network, expanding its potential use-cases.

IOTA IDENTITY & IOTA SMART CONTRACTS

IOTA Identity is a decentralized identity implementation based on the IOTA Tangle. It offers a solution to the problem of data silos and privacy concerns associated with centralized identity systems. It uses the standards developed by the Decentralized Identity Foundation (DIF) and W3C, ensuring interoperability with other Decentralized Identity solutions. Decentralized identities allow individuals, organizations, and even devices to have their own unique, self-sovereign identities that are not controlled by any third-party identity provider. These identities are associated with verifiable, digitally-signed credentials that can prove various attributes or facts about the identity. The key innovation of IOTA Identity is the application of the Tangle, a scalable and feeless DLT, to the field of decentralized identity. This allows identities to be managed in a decentralized way that is both scalable and cost-effective, which could be particularly beneficial for IoT use-cases.

The IOTA Tangle provides a scalable, feeless, and decentralized method of creating and managing identities, and ensures that data privacy and security are maintained at all times. IOTA Identity uses the concept of Decentralized Identifiers, which are globally unique identifiers that do not require a centralized registration authority as they are generated and registered directly on the IOTA Tangle. A DID points to a DID Document, which contains public keys and service endpoints, enabling the DID owner to prove control over the DID and allowing others to find ways to interact securely with the DID. When an identity is created, IOTA Identity publishes the DID Document to the Tangle, which acts as a source of truth for that identity's public keys and service endpoints. IOTA Identity can also handle updates and deletions of identities through the Tangle.

IOTA was in the process of developing a new smart contract platform to be introduced with IOTA 2.0 (Coordicide). IOTA's smart contracts aim to provide the flexibility and functionality of traditional blockchain smart contracts while benefiting from IOTA's scalable, feeless infrastructure. The key innovation here is the adaptation of the smart contract concept to the Tangle. This requires a new approach since the Tangle's structure and consensus mechanism are fundamentally different from traditional blockchains.

Implementing smart contracts in IOTA requires a unique approach because of the nature of the Tangle. Unlike blockchains, the Tangle does not have a linear sequence of blocks, and transactions can be confirmed at different times. IOTA's approach was to introduce "committees" of nodes that reach consensus about the state of a smart contract. Each smart contract has its own committee, and these committees operate independently of each other, allowing for parallel execution of smart contracts. This could potentially push the state of the art in terms of the scalability of smart contracts, as most blockchain-based smart contract platforms have to execute all smart contracts sequentially. A key feature of IOTA Smart Contracts is that they are "flexible" in terms of their security. They can have a large, decentralized committee if they require high security and decentralization, or a small, possibly even single-node committee if they prefer low fees and don't need high decentralization.

EIDAS

eIDAS, which stands for electronic IDentification, Authentication and trust Services, is an EU regulation on electronic identification and trust services for electronic transactions in the internal market. It replaces the Directive 1999/93/EC. The purpose of eIDAS is to provide a regulatory environment to give citizens, businesses, and public administrations in the EU the confidence to carry out digital transactions across borders. The regulation provides standards for digital signatures, digital seals, time stamps, electronic delivery service, and website authentication.

STATE-OF-THE-ART AND CURRENT DEVELOPMENTS

eIDAS has become a widely accepted standard in Europe for digital identity and trust services. It enables enterprises and people to utilize their own national electronic identification schemes (eIDs) to access public services in other EU nations that accept eIDs. The introduction of qualified website authentication certificates (QWACs) and qualified electronic seals is a significant advancement in eIDAS. These characteristics give a high level of confidence for website authentication and assure data integrity and electronic document validity. The eIDAS law also establishes the basis for cross-border digital identities by establishing a legal framework for reciprocal recognition of eIDs among EU member states.

Furthermore, the EU Commission has been working on eIDAS 2.0, which intends to expand the present law to the private sector and cover a larger variety of digital identity and trust services. Another significant step is the EU Commission's announcement in 2021 of a proposed rule on European Digital Identity (EUid). This proposal intends to build a digital identity framework that is generally acknowledged and interoperable across the EU and can be used by people, corporations, and public authorities for a variety of purposes. In terms of blockchain and DLT technologies, as of my previous update, the European Blockchain Services Infrastructure (EBSI) was working on a way to combine eIDAS standards into a blockchain environment. This would enable trustworthy digital identities on blockchain networks, which might pave the way for greater broad use of these technologies in Europe.

The European Blockchain Services Infrastructure (EBSI) is a joint initiative of the European Commission and the European Blockchain Partnership (EBP) to supply cross-border public services throughout the EU using blockchain technology. eIDAS, on the other hand, is an EU rule that establishes the European Union's standard for electronic identity and digital signatures. One of EBSI's primary aims is to use eIDAS to build a foundation for verified digital IDs on its blockchain infrastructure. This would enable for the safe and verified identification of persons, corporations, and entities participating in the blockchain network, which would be critical for many blockchain applications.

While the particular technical specifics of how eIDAS and EBSI are merged may be confidential or in the works, the overall method would most likely include mapping eIDAS digital IDs onto blockchain addresses or accounts. A citizen, for example, may have an eIDAS-compliant digital identity corresponding to a private-public key pair on the blockchain. This manner, any blockchain transactions or interactions they do may be securely linked back to their digital identity, giving the essential degree of confidence and verification. It may also be necessary to store eIDAS-compliant digital signatures on the blockchain. For example, if a citizen digitally signs a document with

their eIDAS digital signature, the transaction might be kept on the blockchain, giving a secure and tamper-proof record of the signing event. To do this, EBSI must guarantee that its blockchain infrastructure can handle eIDAS-compliant digital IDs and signatures. It may be necessary to create particular smart contracts or other on-chain techniques to manage these characteristics.

Envision of TRUSTEE

Incorporating Self-Sovereign Identity, eIDAS, IOTA, and EBSI for cross-border data exchange and reuse throughout the EU would include connecting the capabilities of these systems to build a seamless and secure digital identity and data management infrastructure.

The first step may be to combine SSI and eIDAS to establish a user-centric, highly secure identity management system that adheres to EU requirements. IOTA's decentralized identity architecture might be used to implement SSI, establishing user-controlled and eIDAS-compliant digital identities (DIDs) on the Tangle. The DIDs might then be integrated into EBSI's blockchain services architecture. EBSI may enforce data access and usage guidelines through the use of smart contracts. When data is transferred or a dataset is accessed, a smart contract might validate the user's DID and ensure the user has the required permissions to view or use the data. As part of the authorisation procedure, the user's node's reputation on the IOTA network might be checked (through the Mana system). All transactions involving data exchange and reuse might be immutably recorded on the Tangle for accountability, creating a tamper-proof audit trail. This might be especially important for data provenance and compliance purposes, since it can demonstrate who accessed or utilized a dataset, when, and under what circumstances.

A specific use may be in academic research, where sensitive datasets must be shared and reused across many EU research organizations. A researcher from one institution may utilize their SSI, which is supported by eIDAS and is controlled on the IOTA Tangle, to seek access to an EBSI-administered dataset. The EBSI would use the DID to authenticate the researcher's identity and authorization before granting access to the data. The Tangle may then record the data use, giving an audit trail of the data's use. This multi-layered system, which combines SSI, eIDAS, IOTA, and EBSI, might provide a mix of user autonomy, regulatory compliance, technical efficiency, and security, allowing for the easy and safe exchange and reuse of data across the EU.

OPEN AND FAIR FRAMEWORK FOR EXTRACTED KNOWLEDGE

CONCEPT

An Open and Fair framework can be a central storage for storing various contents such as information, data, knowledge, etc. The framework must provide the functionality where the stored knowledge can be accessed easily, searched and also deleted based on a set time-threshold. It may contain any digital content based on the set procedures, guidelines and information obtained from external components while, at the same time, helping users make decisions to solve problems based on the stored knowledge.

The principal goal of a FAIR framework is to make effective use of the stored content. Hence, the FAIR framework must adhere to data principles to ensure that the stored content is Findable, Accessible, Interoperable and Reusable. An Open and FAIR framework will enhance TRUSTEE to increase the availability of the resources, also keeping in mind that the data is to be generated with longevity; and thus enable the authorised entities to access the stored content to facilitate knowledge discovery.

KNOWLEDGE REPOSITORY AND EXTERNAL TOOLS

A knowledge Repository (KR) often comprises various forms which may include a physical library, a digital database, a content management system and so on. To achieve security, trust and easily accessible functionality, a decentralised framework often plays a crucial role to enhance efficiency, productivity and efficiency. Thus, the Open and Fair framework in TRUSTEE will be implemented by utilising the InterPlanetary File System (IPFS). IPFS is a distributed system which comprises the ability to store, share and retrieve digital content⁶⁵⁹. The IPFS can be integrated with various external components in order to provide enhanced functionality for the stored knowledge;

Monitoring and Orchestration: In order to ensure stability, performance and availability, the KR must be integrated with various powerful tools such as IPFS cluster, IPFS exporter, IPFS private network monitor and so on. Based on the requirements an appropriate tool or combination of tools can be integrated for structured management and optimisation of the infrastructure.

Search Capability: In order to enhance text search capability of the stored contents the KR must be integrated with search engines such as Fleek, temporal and so on. The search capability can enhance the framework to discover contents with index metadata and related contents, a directory of contents, and also enabling access to resources and knowledge sharing efficiently.

Automated Analysis: In order to perform automated analysis on the stored content, the framework can be integrated with a number of automated tools which include Jupyter Notebooks, Apache Spark and on. The data analysis tools can help to obtain insights from the framework. It can also enable statistical analysis and get an extensive overview of the content stored.

Knowledge Graph: In order to avail knowledge graph functionality, the KR should be integrated with various tools. The knowledge graph functionality can provide an overview of the complex data relationship, analysis and inference capabilities.

Blockchain Integration: The combination of IPFS and blockchain technology provides enhanced strengths in terms of decentralisation, data integrity, security, privacy & access control, content indexing, time-stamping and so on. The integration solves many of the drawbacks, thus ensuring scalable data storage solutions. The benefits from the integration of the combined technologies will depend on the specific requirements of the use-case.

IPFS OVERVIEW

The InterPlanetary File System (IPFS)⁶⁶⁰ is a peer-to-peer protocol built by Protocol Labs. IPFS functions similarly to BitTorrent, which is also a peer-to-peer content-sharing protocol. A fundamental difference between the two is that IPFS utilises a content addressing mechanism to locate the stored files whereas BitTorrent or any other web-based protocols make use of a location-based approach to locate a content. One of the main attributes of IPFS is that it does not follow a centralised approach; thus, the stored contents are distributed across various nodes within the IPFS network. The hosted contents can range from databases to websites, media files, documents, apps, AI models and so on. A user can access any content that is hosted in the IPFS network with a content identifier (CID). Being distributed, IPFS have similar functionalities to a blockchain network where several nodes offer their storage bandwidth to store data. The public IPFS network is formed of hundreds of thousands of nodes being open to anyone, whereas the private IPFS provides access to specific groups requiring authorization to join.

IPFS does not use a web address, rather the identifier, which starts with /ipfs, is followed by a content identifier to retrieve a content. An example follows below:

⁶⁵⁹ Y. Psaras, J. M. Soares, D. Dias. 2022. Why the internet need the IPFS. <https://spectrum.ieee.org/peer-to-peer-network>

⁶⁶⁰ IPFS, “What is IPFS?” <https://docs.ipfs.io/concepts/what-is-ipfs/#decentralization>

/ipfs/QmdmQXB2mzChmMeKY47C43LxUdg1NDJ5MWcKMKxDu7RgQm/1004%20-%20Batman

The Content Identifier (CID) is formed on a cryptographic hash of the content enabling the stored content to be retrieved from multiple sources. IPFS utilises Merkle Directed Acyclic Graphs (DAGs) which provide links to the stored content whereas Distributed Hash Tables (DHTs) enable to discover the stored content in the distributed network. The Merkle DAG approach is stimulated by the Git object format which is utilised by the Git protocol for version controlling.

In IPFS, content can be split into blocks where each block can be identified uniquely. This method of splitting blocks enables content to be sourced from several peers. For instance, a 19KB file can be broken into 512 byte chunks where each chunk comprises a unique CID. The main advantage of splitting objects into chunks and then hashing is that when a small update is required then it might not be necessary to update all chunks.

IPFS enables sharing of distributed content without the need for a central server. When compared to the web, the content is obtained by defining the domain name and a path to the resource of the content. (e.g. www.trustee.com/path/myfile.txt). The URL refers to the web-based content, on the other hand, the IPFS neither stores any content centrally nor the stored content is mapped to a particular position. The P2P attribute of IPFS allows to reduce the latency required in sourcing data and creating resilience. The IPFS can be leveraged to store any digital content which may include digital data, images, videos and so on.

Merkle DAG: A MERKLE DAG (Directed Acyclic Graphs) is a DAG which is similar to a MERKLE tree. It ensures that each IPFS node comprises a unique identifier and which is an outcome of the hash of the stored contents. The identifiers are cryptographic hashes generated by algorithms such as SHA-256. The blocks are linked to each other in a graph-like structure which ensures integrity and tamper resistance of the content.

There are some important considerations for MERKLE DAG:

- Merkle DAGS is only formed from the leaves of the nodes without children. The parents get included after the children since it is required to compute the children's identifiers first in order to provide a linkage.
- The Merkle DAG is based on immutable features; thus, changes made to a node would deflect the identifier affecting all connected DAGs.
- All nodes in the Merkle DAG are the root of a sub-Merkle DAG; hence, the subgraph consists of the parent DAG.

IPFS Cluster: An IPFS cluster is a distributed application which permits to have control over the distribution of the content across the nodes within the IPFS network. It allocates and also tracks the pinning of CID across the permitted nodes in an IPFS network. A set of IPFS cluster nodes maintains a set of pinned CIDs, which are also referred to as pin sets. The IPFS cluster node joins to the associated IPFS node thus being liable for retaining the local set of pinned CID of that particular IPFS node. Furthermore, all the IPFS cluster nodes keep up a global pin set having total control by default.

IPFS Pinning: The immutability attributes of IPFS often convince that an entity can simply store contents in IPFS nodes to ensure the immutability of that content. A such misconception often requires further clarification as the immutability in IPFS is referred to as that a CID will also refer to the same content. In addition to that, the time-threshold of contents depends on IPFS nodes that aim to store and share particular contents. In order to maintain a full availability of a content, the associated nodes are responsible for distributing the content to their peers.

The IPFS comprise a default functionality in which the nodes cache the content they request from other nodes and also share it when required. In order to stop contents from being shared within the cache, the policy required to be changed as below:

```
1
>
ipfs
```

config Reprovider . Strategy pinned

Due to the limitations in the storage capability, the nodes use a garbage collection policy which enables data deletion from the cache when not used. However, in order to restrict data deletion, the pinning concept⁶⁶¹ should be exercised. The contents in a node which has been pinned do not get deleted from the garbage collection process.

REVIEW OF THE USE-CASES

Chen et al.⁶⁶² propose an improved P2P file system scheme integrating IPFS and Blockchain. The authors have identified the high-throughput challenges that individual users face within IPFS. A unique zigzag-based storage model is proposed which not only is able to refine the block storage model of IPFS but also helps overcome issues concerned with data reliability and availability, storage overhead and service providers.

Muralidharan et al.⁶⁶³ provide an approach which is implemented on an IoT network utilising IPFS. In their approach, the content is addressed using unique cryptographic hashes. A DHT enables one to look at and also publish data for peers. Furthermore, the Merkle DAG along with Git versioning keeps the distributed network up-to-date. In order to ensure security, their proposed method leverages secure file sharing and encrypted communication.

Zheng et al.⁶⁶⁴ proposed an innovative method which combined IPFS and blockchain. In their method, the miner puts the transaction data into the IPFS and stores the hash of the transaction in the blockchain. This enables reducing the blockchain data. According to their experiment, the compression ratio may get to 0.0817 and thus the analysis defines it to be a good performance in terms of security and synchronisation.

Steichen et al.⁶⁶⁵ propose a blockchain-based decentralised access control for IPFS. The authors have presented a transformed version of IPFS which utilises Ethereum smart contracts in order to allow access-controlled file sharing. The main aim of the smart contract is to arrange the access control list, while the modified IPFS application executes it. In order to achieve this, the smart contract is interacted with right after content is uploaded, downloaded or transferred.

Kumar et al.⁶⁶⁶ propose a distributed storage for healthcare systems utilising blockchain and IPFS. They have identified the security issues which include the unauthorised access to users confidential information such as identity

⁶⁶¹ Pinata.cloud, "Pinata — Effortless IPFS File Management," 2021. Available: <https://pinata.cloud/>

⁶⁶² Y. Chen, H. Li, K. Li and J. Zhang, "An improved P2P file system scheme based on IPFS and Blockchain," 2017 *IEEE International Conference on Big Data (Big Data)*, Boston, MA, USA, 2017, pp. 2652-2657, doi: 10.1109/BigData.2017.8258226.

⁶⁶³ S. Muralidharan and H. Ko, "An InterPlanetary File System (IPFS) based IoT framework," 2019 *IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2019, pp. 1-2, doi: 10.1109/ICCE.2019.8662002.

⁶⁶⁴ Q. Zheng, Y. Li, P. Chen and X. Dong, "An Innovative IPFS-Based Storage Model for Blockchain," 2018 *IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Santiago, Chile, 2018, pp. 704-708, doi: 10.1109/WI.2018.000-8.

⁶⁶⁵ M. Steichen, B. Fiz, R. Norvill, W. Shbair and R. State, "Blockchain-Based, Decentralized Access Control for IPFS," 2018 *IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Halifax, NS, Canada, 2018, pp. 1499-1506, doi: 10.1109/Cybermatics_2018.2018.00253.

⁶⁶⁶ R. Kumar, N. Marchang and R. Tripathi, "Distributed Off-Chain Storage of Patient Diagnostic Reports in Healthcare System Using IPFS and Blockchain," 2020 *International Conference on COMMunication Systems & NETworks (COMSNETS)*, Bengaluru, India, 2020, pp. 1-5, doi: 10.1109/COMSNETS48256.2020.9027313.

information, confidential medical information and so on. In order to address the challenges, they have proposed the distributed off-chain method. The proposed method will ensure privacy by providing access only to authorised entities and at the same time it will achieve consistency, integrity and data availability.

Hao et al.⁶⁶⁷ propose an efficient Storage Scheme for agricultural product tracking. The authors have leveraged a distributed data storage model utilising IPFS and blockchain. In their method, IPFS is used to store digital contents such as video, images, and monitored data from the sensor. In order to enhance security and privacy the blockchain is adopted to record the hash of the stored content. To achieve that, they have designed a blockchain-based authentication mechanism which not only verifies the data but also ensures proper data security. Their experiment evidences that the proposed method is efficient compared to existing methods.

PRIVACY, ANTI-CENSORSHIP AND IMMUTABILITY

The IPFS network can be based on both public and private methods. One of the main differences between the private and public IPFS is that in a private IPFS network, only the permissioned nodes are permitted to join the network, thus the network can only scale to those nodes. When providing access, a more careful and access control approach should be adopted as giving access to unauthorised parties can result in both security and privacy being compromised.

The public IPFS utilises publicly available consistent content addresses which are advertised through publicly viewable DHTs. Thus, the public IPFS does not contain a privacy layer but can be achieved by overlaying on top, for instance, by encrypting private information, forming a private IPFS network and so on. One of the main attributes of public IPFS is that it can alleviate the option to censor content. Due to the P2P nature of the IPFS protocol, as long as a single peer remains with the hosted content, a user will be able to retrieve the required content hosted on it. However, the freedom of sharing, in a Public IPFS, without any censorship in place, can often cause several challenges when illegal content is shared in the network alongside legitimate content.

Furthermore, the immutability nature can also cause some challenges as IPFS does not support automated data deletion facilities. Juan Benet, the creator of IPFS, points out that "Objects are permanent"; thus, any content that is added to IPFS remains permanent whereas immutability is a design by choice in order to obtain high performance and data integrity.

LIMITATIONS

The privacy, distribution of illegal content, and immutability of data are not the only concerns in a public IPFS, many practical considerations should also be considered when implementing public IPFS. Some of the limitations are as follows;

Lack of Metadata: It can often be essential to obtain metadata such as timestamps, upload time, creator of content, author and so on. However, such information is not guaranteed to exist in files that are shared using the IPFS protocol. Every content shared in the network gets identified by CID and thus the absence of such information can be very essential when creating a powerful audit trail mechanism.

Update Control: Besides the lack of version-controlling features, IPFS also does not have the capability to ensure whether a user has retrieved the latest version of the uploaded content. For instance, when a content is updated, a new CID is generated to refer to that new version. However, the original version of the uploaded content still remains accessible with the associated CID. As a result, the newly generated CID is required to be shared with all nodes to notify them about the updated content. IPFS also provides a service which is referred to as InterPlanetary

⁶⁶⁷ Hao, J., Sun, Y. and Luo, H., 2018. A safe and efficient storage scheme based on blockchain and IPFS for agricultural products tracking. *Journal of Computers*, 29(6), pp.158-167.

Name System (IPNS)⁶⁶⁸ can be utilised to map content in order to provide a fixed address for the uploaded content. However, IPNS may not be suited well where an entity aims to have full control over the content they prefer to obtain.

Content Indexing: An entity obtains content utilising the identifier or the content; thus, the user must have prior knowledge about the CID before they obtain the content. Hence, IPFS must have the search capability of the content and in order to implement such functionality the shared content, index of the content as well as the CID may require to be implemented manually.

PRIVATE IPFS

In a private IPFS, content sharing occurs among trusted nodes in a private network. There are various techniques which can be utilised to achieve the permissioned network such as access control mechanisms, etc. A private IPFS network is generally useful when it is essential to store sensitive data where data sharing is required to be restricted to certain entities. One of the major advantages of a private IPFS is that it ensures enhanced privacy and control over the stored content. Furthermore, the access control enables increased security and thus safeguards content from unauthorised access, malicious attacks and so on. Similarly, a private IPFS provides more control over the stored content; such as content deletion option, etc. In addition to that, performance plays a significant role in private IPFS. Since a private IPFS network involves a smaller number of trusted nodes, the entire process from file retrieval to distribution occurs at a fast pace, especially as it makes the process more reliable when processing large or sensitive content.

Advantages of Private IPFS

The private IPFS provides several security advantages over public IPFS.

- Access control: Restricted access ensures that only the authorised entities have control over the stored knowledge. Access control helps mitigate unauthorised access by malicious entities to carry out exploitation.
- Confidentiality: Full confidentiality can be ensured in a private IPFS while sharing and storing sensitive information. Unlike public IPFS, the disclosure of sensitive data does not occur and remains within the trusted network.
- Enhanced Security methods: Private IPFS offer this functionality to implement customer security methods based on the requirements. Thus, such measures can be aligned based on the consortium's security policies, standards and so on.
- Reduced Attack Surface: The nodes are the potential entry points of attacks. In a private IPFS, the limited number of nodes reduces the vulnerabilities an attacker could exploit to obtain information.
- Incident response: In a private IPFS the incident response process can be faster by integrating various threat detection and visualisation tools. Furthermore, the controlled nature of the infrastructure helps respond to incidents at a fast pace.

XAI TECHNIQUES FOCUSING ON TRANSPARENCY, TRUSTWORTHINESS AND FEDERATED ML

EXPLAINABLE AI OVERVIEW

Nowadays, AI has been gaining significant attention as it advances towards practical applications in various domains. AI algorithms can help to address challenges that were impossible to solve in the past years. However, the increasing performance of these algorithms has led to a consequent increase in their complexity. It is not

⁶⁶⁸ Protocol Labs, “IPNS — IPFS Docs,” 2021. Available:<https://docs.ipfs.io/concepts/ipns/#example-ipns-setup-with-cli>

uncommon for many ML models that are deployed in production to be considered as black-box models, with their internal workings and decision making process to be difficult to be understood even by experts on this field.

The lack of understanding that is accompanied with AI algorithms can lead potentially to a lack of trust, which in turn can adversely affect the application scale of these technologies. In critical domains, like in healthcare or in the automotive industry, the decisions of AI models can affect the society to a large extent. Furthermore, European Union's introduced General Data Protection Regulation (GDPR), creates a right to explanation, which means that a user can require the explanation of a algorithmic model decision⁶⁶⁹. As a result, it is critical for the involved stakeholders to be able to understand on which grounds the AI model makes a prediction, and what is the relationship between the input data and the output. That way, they can trust more this model and further extend the range of applications.

Explainable Artificial Intelligence (XAI) techniques have been developed as a solution to this problem. In general, XAI attempts to develop methods that can provide to stakeholders and end-users explanations about the decision-making process of AI systems. This can result in enhancing the transparency of black-box models and increasing the trust on such applications. Especially in industrial applications, sometimes models that have poorer performance, but are interpretable, might be preferred⁶⁷⁰. XAI techniques can assist to the trade-off between interpretability and performance/complexity, by providing ways to explain the decision making process of accurate but complex black-box models⁶⁷¹.

The main objective of XAI is to improve the Transparency and Understanding of AI models. To achieve this, XAI techniques provide information about the decision-making process of the model, as well as the problem that is addressed. AI models are often complex and in many cases it can be challenging for humans to understand why a model made a particular decision, especially if it utilizes black-box algorithms. XAI attempts to improve transparency by providing human-understandable information in the form of visualizations or natural language explanations⁶⁷².

Furthermore, another objective of XAI is to improve Trustworthiness. Trust is an essential factor for the widespread adoption and use of AI, especially in critical domains where the AI decisions can have a significant impact⁶⁷³. When the involved stakeholders can get an understanding of the reasoning behind the AI model decisions, then it is most likely to accept and use these models.

As AI is more and more involved in social environments, the explainability of these systems aims to establish Fairness. In certain occasions, there is the possibility that an AI model can be trained on biased data, which can

⁶⁶⁹ B. Goodman and S. Flaxman, "European union regulations on algorithmic decision-making and a "right to explanation"," AI Magazine, vol. 38, no. 3, pp. 50–57, oct 2017. [Online]. Available: <https://doi.org/10.1609/v38i3.2741>

⁶⁷⁰ L. Veiber, K. Allix, Y. Arslan, T. F. Bissyandé, and J. Klein, "Challenges towards Production-Ready explainable machine learning," in 2020 USENIX Conference on Operational Machine Learning (OpML 20). USENIX Association, Jul. 2020. [Online]. Available: <https://www.usenix.org/conference/opml20/presentation/veibe>

⁶⁷¹ A. B. Arrieta, N. D. Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, Daniel Molina, R. Benjamins, R. Chatila, F. Herrera, "Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI," CoRR, vol. abs/1910.10045, 2019. [Online]. Available: <http://arxiv.org/abs/1910.10045>

⁶⁷² A. Weller, "Challenges for transparency," CoRR, vol. abs/1708.01870, 2017. [Online]. Available: <http://arxiv.org/abs/1708.0187>

⁶⁷³ F. K. Došilović, M. Brčić, N. Hlupić, "Explainable artificial intelligence: A survey," in 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2018, pp. 0210–0215

result in discriminatory results⁶⁷⁴. By understanding how the model utilizes the input data and uncovering the relations that lead to a certain result, XAI can highlight any biased use of data and unfair decisions. In this sense, XAI aims to prevent the unethical and unfair use of AI models.

Another objective of XAI is to ensure Accountability. As AI-generated decisions become increasingly important, it is crucial to be able to trace how the system arrived at a particular decision. XAI can help to ensure that AI models are reliable, meaning that they can perform consistently and accurately over time⁶⁷⁵. In addition, XAI can assist with the establishment of guidelines and policies to regulate the use of AI.

Algorithms used in AI can be categorized as transparent and black-box. The results and the process with which these are obtained by transparent models can be understood by analyzing the model internal structure. However, in practical applications most AI models are not transparent by design, either because their architecture is not specified and made publicly available, or because they are too complicated. These models are considered as black-box models even by AI experts, who can not fully understand the decision-making process. In this case, post-hoc XAI techniques can be applied for the explanation of models that are not transparent by design.

XAI APPLICATIONS IN THE ENERGY

XAI tools for solar photovoltaic power generation forecasting have appeared in the relevant literature⁶⁷⁶. The power generation forecasting is formed as a regression problem, and the XGBoost algorithm is utilized in this approach. The dataset consists of real measurements from a PV station, and includes the PV power output, weather conditions and time index, collected in 15-minute intervals over a period of 6 months. The ELI5 tool is used to interpret the XGBoost algorithm, and a feature importance analysis is performed both on all the features and on a subset of the features in order to find the most important ones. Individual predictions are also explained by ELI5, showing how each feature contributes to the prediction. The algorithm is trained and evaluated both on all the features, and on a subset of the features and the RMSE score is used as an evaluation metric. The results show that by using all the features the model achieves a lower RMSE score and better predictions.

A framework based on XAI and ML classifiers in order to investigate explainability on energy building performances has also been introduced in the literature⁶⁷⁷. After pre-processing the data, several classifications algorithms, including Decision Trees (DT), Random Forest (RF), Extra Trees (ET), Bagging classifiers (BC) and Multilayer Perceptrons (MLP) are used to assign the flats to different energy performance classes. The k-means clustering algorithm is used next, in order to identify the most relevant predictions of the correct and wrong classified instances. The LIME XAI algorithm is used to explain these instances and find the impact of the features on these predictions. The framework is applied on a dataset consisting of Energy Performance Certificates (EPCs) for 100000 flats located in a region of Italy. There are four main categories of features in the dataset, relevant to

⁶⁷⁴ M. Kim, O. Reingold, and G. Rothblum, "Fairness through computationally-bounded awareness," in *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., vol. 31. Curran Associates, Inc., 2018. [Online]. Available: <https://proceedings.neurips.cc/paper/2018/file/c8dfece5cc68249206e4690fc4737a8d-Paper.pdf>

⁶⁷⁵ M. Khan and J. Vice, "Toward Accountable and Explainable Artificial Intelligence Part one: Theory and Examples," 2022. [Online]. Available: [https://www.techrxiv.org/articles/preprint/Toward Accountable and Explainable Artificial Intelligence Part one Theory and Examples/19102085](https://www.techrxiv.org/articles/preprint/Toward_Accountable_and_Explainable_Artificial_Intelligence_Part_one_Theory_and_Examples/19102085)

⁶⁷⁶ S. Sarp, M. Kuzlu, U. Cali, O. Elma, and O. Guler, "An interpretable solar photovoltaic power generation forecasting approach using an explainable artificial intelligence tool," in *2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2021, pp. 1–5.

⁶⁷⁷ A. Galli, M. S. Piscitelli, V. Moscato, and A. Capozzoli, "Bridging the gap between complexity and interpretability of a data analytics-based process for benchmarking energy performance of buildings," *Expert Systems with Applications*, vol. 206, p. 117649, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417422009526>

geometry, envelope, time and system. The results highlight that the proposed method can break the accuracy and interpretability trade-off that usually limits the energy performance analysis, can better refine the model, and explain which features influence strongly the energy performance of a building.

XAI APPLICATIONS IN AUTOMOTIVE

An XAI method⁶⁷⁸ for CNN models utilized by self-driving cars builds upon an algorithm that focuses on Sensitivity Analysis, which can estimate the influence of each input feature. First, the images are converted to gray-scale and are modified by making changes from the original image. Then, the original and modified images are input to the network and their difference is calculated by comparing the output vectors. Then, an explanation is made by feeling each part of the original image proportionally to the previously calculated difference. The explanation highlights the part of the image that influences the final prediction. The XAI method is showcased with a pre-trained CNN with four image datasets, that include vehicle/non-vehicle images and images with traffic signs. Individual predictions are explained by this method, highlighting each step of the algorithm, and the explanations are compared to popular XAI methods SHAP, LIME, Grad-CAM and XCNN. The results show that the proposed method does not fail for any image, unlike the other methods that can fail for some images and provide unsuccessful explanations

An XAI method for trustworthy In-Vehicle Intrusion Detection System (IV-IDS) has also appeared in the literature with the aim to provide explainable predictions for the detection of cyber-attacks on automotive vehicles⁶⁷⁹. The detection system consists of a DNN with three hidden layers and is trained and evaluated using a dataset that contains 1735840 instances of which 183314 are attacks. The dataset is split in 70% for training and 30% for testing, and the model achieves a 99.73% accuracy after a 10-fold cross validation on the training dataset. To explain the predictions, a visualization-based tool called VisExp is designed and implemented, that aggregates local SHAP explanations to provide insight about the global behaviour of the DNN model. To evaluate the effectiveness of the method, a survey is created and distributed to domain experts in the fields of AI, cybersecurity and automotive. An analysis is performed on the survey results, which highlights that the proposed tool can enhance the explainability and trustworthiness of automotive intrusion detection systems.

XAI APPLICATIONS IN HEALTH

An explainable machine learning model⁶⁸⁰ for the early detection of Parkinson's disease (PD) has been proposed in the relevant literature. More specifically, they worked on the medical imaging task of classifying a given SPECT DaTscan as having Parkinson's disease or not. The dataset came from the Parkinson's Progression Markers Initiative (PPMI) database and included 642 DaTscan SPECT images from the initial screening of unique patients. Every SPECT DaTscan was initially presented as a 3D volume space in DICOM and NIFTI format and after visual analysis and based on previous studies authors chose the slice that was extracted and converted to jpeg format. They applied several preprocessing techniques including cropping, augmentations and change of size. Data were then trained on a deep neural network using transfer learning. They used Convolutional Neural Networks (CNN) and more specifically the VGG16 neural network architecture in which they replaced the final layer for task classification and retrained the last two CNN layers. Different metrics were used to measure the performance like accuracy,

⁶⁷⁸ H.-S. Kim and I. Joe, "An xai method for convolutional neural networks in self-driving cars," PLOS ONE, vol. 17, no. 8, pp. 1–17, 08 2022. [Online]. Available: <https://doi.org/10.1371/journal.pone.0267282>

⁶⁷⁹ H. Lundberg, N. I. Mowla, S. F. Abedin, K. Thar, A. Mahmood, M. Gidlund, and S. Raza, "Experimental analysis of trustworthy in-vehicle intrusion detection system using explainable artificial intelligence (xai)," IEEE Access, vol. 10, pp. 102 831–102 841, 2022.

⁶⁸⁰ P. R. Magesh, R. D. Myloth, and R. J. Tom, "An Explainable Machine Learning Model for Early Detection of Parkinson's Disease using LIME on DaTSCAN Imagery," Computers in Biology and Medicine, vol. 126, p. 104041, Nov. 2020. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0010482520303723>

specificity, sensitivity, precision, Cohen's Kappa score, F1 score, and area under the ROC curve. For the classification between PD and non-PD, they used a threshold to normalize the predicted probabilities. They analysed the curve to understand the trade-off between the TPR and FPR for different thresholds and chose the one that provided the best performance. Explanations were given using the visual indicators generated by the LIME method. The LIME explainer attributed the superpixel markings highlighting the ROI in the data as the parts of the image that had the most influence in determining the outcome of the prediction. LIME ultimately allowed the visual tracing of the ROI, which helped non-experts understand the diagnosis of the patient.

An analysis of various explainable machine learning prediction models for hospital mortality can also be found in the XAI literature for the health domain⁶⁸¹. They developed various ML models and examined their internal behaviour with SHapley Additive exPlanations (SHAP) values. The exact classification task was to predict the outcome of patients in the ICU. They used the publicly available eICU database, which contained information about patients admitted to critical care units in the US such as patient demographics, vital values, lab values, and diagnoses. The selected features were the same as for calculating the APACHE IV score with some exceptions. The four ML models that the authors developed were Random Forest (RF), Logistic Regression (LR), Adaptive Boost Classifier (ADA), and Naive Bayes (NB). Another goal of this study was to inspect how individual features impacted predictions from different ML models and find similarities and differences between the models. Regarding the output from the SHAP method, authors produced SHAP summary plots which showed the feature importance for the different ML models, SHAP force plots that depicted the contribution of an individual feature for one or several patients and lastly and individual force plots which included the features that each model considered relevant for the prediction of individual patients. Regarding the model comparison, the RF model had the highest AUC, followed by the ADA model, APACHE IV, LR and NB. They found that the ML models had comparative discriminative abilities, and similar feature importance values but had different behaviour. Another conclusion was that there were differences between the factors that were deemed important by the models and the ones based on medical theory.

XAI APPLICATIONS IN EDUCATION

An explainable model for the identification of higher education students that were at risk of failing the preparatory year at an early stage has appeared recently in the XAI literature for the education domain⁶⁸². The study included the application of different data mining algorithms on educational data and the utilization of XAI techniques to explain the proposed prediction models. The datasets used in this study contained student information related to their preparatory year from a single Saudi Arabian university (IAU). Data came from various sources like a Student Information System and a Learning Management System and included demographic data, pre-university data, course grades, Cumulative Grade Points Average (CGPA), and assessment marks for various courses. The specific task was to classify the CGPA level of a preparatory-year student. For this reason, the authors used the students' first-year courses and grades to construct three classification models. The first model was created to predict whether a student was at risk of failing the preparatory year or not and to find the most influential features. After finding the highest influence courses, two additional models were built to predict the performance of a student in the two courses of the highest influence. Authors implemented and evaluated different supervised learning methods including Random Forests (RF), Support Vector Machines (SVM), and Artificial Neural Networks (ANN). The authors applied explainable AI techniques like the Local Interpretable Model-agnostic Explanations method (LIME), SHapley Additive exPlanations (SHAP), and the global surrogate model to explain the predictions of the

⁶⁸¹ E. Stenwig, G. Salvi, P. S. Rossi, and N. K. Skjærvold, "Comparative analysis of explainable machine learning prediction models for hospital mortality," *BMC Medical Research Methodology*, vol. 22, no. 1, p. 53, Dec. 2022. [Online]. Available: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-022-01540-w>

⁶⁸² S. Alwarthan, N. Aslam, and I. U. Khan, "An Explainable Model for Identifying At-Risk Student at Higher Education," *IEEE Access*, pp. 1–1, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9906532/>

black-box models. LIME provided a visual explanation for Pass and Fail data samples including the prediction probabilities, the relative significance and the actual value for the most significant features. SHAP was able to explain predictions for both a single data sample and the entire dataset. A summary plot depicted the global contribution of features for the two courses. Lastly, the authors used surrogate decision tree models that approximated the complicated RF model and displayed the first three layers of the model for the two courses. They tested and evaluated different pre-processing techniques. They used two sampling approaches SMOTE, and SMOTE-Tomek Link to handle the class imbalance and three feature selection methods including Recursive Feature Elimination (RFE) with Pearson correlation coefficient, RFE with mutual information, and Genetic Algorithm (GA) to find the most relevant features for building the final model. For the evaluation authors used grid and randomized search with 5-fold cross-validation and accuracy, precision, recall, and F1 score as performance metrics. The results showed that RF was the highest-performance model followed by SVM and ANN.

An explainable approach for the prediction of students' performance has also been proposed⁶⁸³. The goal was to identify the students that had learning difficulties or were likely to fail a course. The authors used data collected from a learning management system (LMS) and contained information about the student's involvement in learning and teaching. They developed both global and local models and applied XAI methods to analyze the learning process of each student as well as to find the most important factors that affected their learning effectiveness. For the model performance, the authors used metrics like accuracy, Kappa, sensitivity and specificity that were derived from a confusion matrix. The implemented system included three stages. In the first stage, they used decision tree (DT), deep neural network (DNN) and K-nearest neighbours (KNN) as global models for the prediction of the learning outcome. The DNN model achieved the highest accuracy while the DT model was the chosen global model due to its explainability. In the second stage, they used a model grafting technique by combining a global DT model and a local DNN model (DT-DNN). This hybrid model allowed users to achieve the desired balance between accuracy and interpretability. The prediction accuracy of the DT-DNN system slightly increased when compared to the original DT model but it failed to outperform the DNN model. The authors used LIME for the interpretation of the DNN model and pinpointed the most important features. They also interpreted the other two global models and produced supplementary results of important features. Furthermore, the authors implemented a compensatory and a non-compensatory approach to increase the explainability of the hybrid model. The multiple antecedent statements of rules derived from the DT model were combined with the variable importance derived from DNN(LIME) model. The compensatory approach considered the importance of duplicated variables derived from both models while the non-compensatory approach compared the duplicated variables and ignored the variable importance from the local models. In the third stage, they used another model fusion technique where they integrated a DT model with a KNN model as a single global model. Here, they merged the decision rules or variable importance with the output of the hybrid model. This model could obtain a prediction without using a proxy model like LIME to provide explainability. According to the results, this technique reduced the type I and II errors.

XAI APPLICATIONS IN SPACE

The use of XAI for the detection of compact stellar systems and more specifically Ultra-Compact Dwarfs (UCDs) and Globular Clusters (GCs) around galaxies has been recently considered in the space domain⁶⁸⁴. This was a multi-class classification problem where the goal was to separate UCDs and GCs from the foreground stars and background galaxies. For this task, they used multi-wavelength imaging data of the Fornax galaxy cluster and optical and near-infrared measurements. The dataset consisted of 23 features with each observation belonging to a class which could be background galaxies, UCDs and GCs or nearby foreground stars. For balancing the data they

⁶⁸³ T. Chou, "Apply explainable AI to sustain the assessment of learning effectiveness," in 12th International Multi-Conference on Complexity, Informatics and Cybernetics, IMCIC 2021, 2021, pp. 113–118

⁶⁸⁴ M. Mohammadi, J. Mutatiina, T. Saifollahi, K. Bunte, "Detection of extragalactic Ultra-compact dwarfs and Globular Clusters using Explainable AI techniques", in *Astronomy and Computing*, Volume 39, 2022

used Synthetic Minority Oversampling (SMOTE) and Borderline SMOTE. The authors employed two machine learning algorithms, namely Localized Generalized Matrix Learning Vector Quantization (LGMLVQ) and Random Forest (RF). The authors compared the models' performance and feature importance. For the evaluation of the classification performance, the authors used the training and test errors with their standard deviations, binary class measures such as purity and completeness and the confusion matrix. According to their results, the RF and LVQ models achieved comparable performance.

XAI BY DESIGN

In the literature of the inverse problems such as denoising, deblurring and super-resolution, it is quite common to formulate optimization problems that aim to capture and incorporate information regarding the underlying physical properties of the considered problem. Considering the highly ill-posed nature of these problems, several approaches utilize proper priors, referred to as regularization terms, thus forming a cost function as follows.

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \left(\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_F^2 + \lambda \cdot \mathbf{R}(\mathbf{x}) \right)$$

where $\mathbf{x}^*$ denotes the estimated signal, $\mathbf{H}$ is the forward model such as blurring kernel and $\mathbf{R}(\cdot)$ stands for the regularization term. The role of such terms in the cost function is to “promote” (i.e., give less cost) to solutions that possess “favorable” characteristics. Thus, a significant body of research has tried to devise proper mathematical hand-crafted regularization terms (e.g., sparsity norms, nuclear norm, total-variation) that capture such favorable characteristics, for various inverse problems. Although the hand-crafted regularizers provide satisfactory performance results in a wide range of applications, the fact that they are quite general poses some limitations, since they are not designed to capture specific properties from the signals of interest.

In recent years, an emerging body of studies have explored the use of regularization terms that are properly learnt from suitable training data to capture the dependencies of the signals existing on the examined application^{685,686,687}. These regularization terms are used in place of the “hand-crafted” regularization terms traditionally employed in the cost functions of the studied restoration problems and have the form of a suitable artificial neural network. Appealing models for “such” regularizers constitute the convolutional neural networks, due to their powerful representation capacity and the advantage of GPU computation.

Having formed the optimization problem, an efficient iterative solution can be derived utilizing some iterative optimization solver e.g., Alternating Direction Method of Multipliers⁶⁸⁸ (ADMM) or Half Quadratic Splitting⁶⁸⁹ (HQS). In light of this, a simple strategy, known as “plug-and-play” method⁶⁹⁰, can be used where the neural network is optimized from proper training data, and then it is plugged into the optimization solver. This procedure

⁶⁸⁵ Gilton, G. Ongie, and R. Willett, “Learned patch-based regularization for inverse problems in imaging,” in 2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2019, pp. 211–215.

⁶⁸⁶ Zhang, W. Zuo, S. Gu, and L. Zhang, “Learning deep cnn denoiser prior for image restoration,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.

⁶⁸⁷ G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, “Deep learning techniques for inverse problems in imaging,” IEEE Journal on Selected Areas in Information Theory, vol. 1, no. 1, pp. 39–56, 2020.

⁶⁸⁸ Boyd, Stephen, et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers." Foundations and Trends® in Machine learning 3.1 (2011): 1-122.

⁶⁸⁹ Geman, Donald, and Chengda Yang. "Nonlinear image recovery with half-quadratic regularization." IEEE transactions on Image Processing 4.7 (1995): 932-946.

⁶⁹⁰ Zhang, W. Zuo, S. Gu, and L. Zhang, “Learning deep cnn denoiser prior for image restoration,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.

is repeated until convergence is reached. However, we should emphasize that the “plug-and-play” is sub-optimal, has a limitation, since the neural network (regularizer) is trained independently of the inverse problem at hand and the degraded data, thus requiring massive amounts of training data and a large number of iterations to obtain satisfactory results⁶⁹¹.

To address the limitations of the plug-and-play approaches, more efficient and effective methodologies have been proposed based on the model-based deep learning theory, known as deep unrolling methods^{692,693,694,695,696,697}. Having formed a proper iterative solver as mentioned above, the deep-unrolling strategy aims to transform the well-justified optimization solver into a meaningful deep learning architecture. In more detail, by unrolling a small number of iterations of the iterative optimization problem, a deep network is derived, where each layer of this architecture corresponds to a unique iteration of the solver. Thus, if we unroll the solver for K iterations, a K -layer deep learning network is formed, where its features and architecture such as number of layers and parameters are highly interpretable and well justified. Figure 19 illustrates the derived architecture.

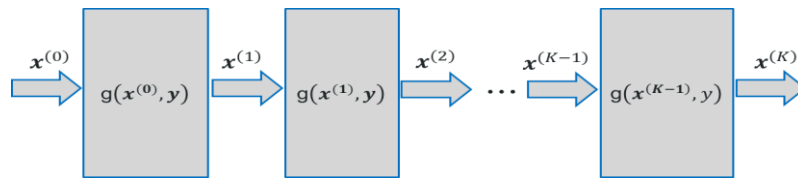


Figure 19: Having form an iterative solver for the optimization problem in (1) , denoted as $g(x, y)$, a deep learning architecture is derived where each layer of this network corresponds to a unique iteration of the solver.

Adopting a wider perspective, the deep-unrolling model-based methods provide numerous significant benefits in various real-world problems and applications. In particular, compared the classical deep learning approaches where their architectures derive from an ad-hoc trial-and-error manner, these deep unrolling networks have a clear connection with the under-examined problem and the considered optimization algorithm⁶⁹⁸. Along with their high interpretability, the derived deep unrolling models are more concise, requiring less training data and number of parameters to achieve state-of-the-art performance compared to the classical deep learning networks. This stems from the fact that the unrolling models consider the underlying physical processes and utilize prior domain knowledge of the examined problem.

LESSONS LEARNT

⁶⁹¹ . Gilton, G. Ongie, and R. Willett, “Deep equilibrium architectures for inverse problems in imaging,” IEEE Transactions on Computational Imaging, vol. 7, pp. 1123–1133, 2021.

⁶⁹² . Monga, Y. Li, and Y. C. Eldar, “Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing,” IEEE Signal Processing Magazine, vol. 38, no. 2, pp. 18–44, 2021.

⁶⁹³ Yang, J. Sun, H. Li, and Z. Xu, “Admm-csnet: A deep learning approach for image compressive sensing,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 3, pp. 521–538, 2020.

⁶⁹⁴ Y. Chun, Z. Huang, H. Lim, and J. Fessler, “Momentum-net:Fast and convergent iterative neural network for inverse problems,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1–1, 2020

⁶⁹⁵ . Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar, “Efficient and interpretable deep blind image deblurring via algorithm unrolling,” IEEE Transactions on Computational Imaging, vol. 6, pp. 666–681, 2020.

⁶⁹⁶ . Gkillas, D. Ampeliotis, and K. Berberidis, “Connections between deep equilibrium and sparse representation models with application to hyperspectral image denoising,” IEEE Transactions on Image Processing, 2023.

⁶⁹⁷ M. Scetbon, M. Elad, and P. Milanfar, “Deep k-svd denoising,” IEEE Transactions on Image Processing, vol. 30, pp. 5944–5955, 2021.

⁶⁹⁸ M. Scetbon, M. Elad, and P. Milanfar, “Deep k-svd denoising,” IEEE Transactions on Image Processing, vol. 30, pp. 5944–5955, 2021.

The detailed literature review presented above explores the use of XAI techniques in six critical domains, where the use of AI is continuously advancing; Energy, Automotive, Health, Education, and Space. XAI applications in Energy can help to identify which features affect mostly power generation forecasting or building energy consumption. In Automotive, XAI methods can increase the transparency of the computer vision models that are commonly used for self-driving and explain how these models utilize road and traffic signs information. XAI applied in Healthcare can help to increase trust on AI based diagnostic tools and explain why the model reached a certain decision-diagnosis. The application of XAI in education can make sure that AI models meet certain regulations, related to fairness and ethics, that are essential in this particular domain. Furthermore, XAI techniques can be applied to further utilize the advantages of AI in space exploration. To conclude, XAI applications have a promising future in these domains, but further research and development is needed in order to address challenges such as scalability of these methods and evaluation of the explanations.

PRIVACY AND SECURITY IN FEDERATED LEARNING: CHALLENGES, SOLUTIONS AND APPLICATIONS

In today's data-driven world, most applications and services such as healthcare and medical services^{699,700}, natural language processing⁷⁰¹, autonomous vehicles⁷⁰², smart manufacturing⁷⁰³, recommender systems⁷⁰⁴, geophysical exploration⁷⁰⁵, to name a few, are based on AI technology with complex data-hungry ML algorithms, usually involving the use of Deep Neural Networks (DNN)⁷⁰⁶. In traditional ML algorithms, centralized data processing and management systems are inherently prone to a single point-of-failure (including the risk of severe data breaches) which could lead to the lack of trust from end-users well as the difficulty in complying with the GDPR⁷⁰⁷. At the same time, the emergence of centralized searchable repositories has turned the leakage of private information (including health, travel, and financial data) into an urgent social problem⁷⁰⁸, leading to growing concerns for data privacy among the involved parties and rendering privacy preservation an issue of critical importance⁷⁰⁹.

In recent years, federated learning (FL)⁷¹⁰ has emerged as a new ML paradigm for achieving collaborative training and data privacy at the same time, by decoupling the provision of data at the end-user from the ML model

⁶⁹⁹ Cao, C., Liu, F., Tan, H., Song, D., Shu, W., Li, W., ... & Xie, Z. (2018). Deep learning and its applications in biomedicine. *Genomics, proteomics & bioinformatics*, 16(1), 17-32.

⁷⁰⁰ Suzuki, K. (2017). Overview of deep learning in medical imaging. *Radiological physics and technology*, 10(3), 257-273.

⁷⁰¹ Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, 13(3), 55-75.

⁷⁰² Xu, Y., Yang, X., Gong, L., Lin, H. C., Wu, T. Y., Li, Y., & Vasconcelos, N. (2020). Explainable object-induced action decision for autonomous vehicles. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9523-9532).

⁷⁰³ Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of manufacturing systems*, 48, 144-156.

⁷⁰⁴ Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)*, 52(1), 1-38.

⁷⁰⁵ Yu, S., & Ma, J. (2021). Deep learning for geophysics: Current and future trends. *Reviews of Geophysics*, 59(3), e2021RG000742.

⁷⁰⁶ Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016). <http://www.deeplearningbook.org>

⁷⁰⁷ Truong, N. B., Sun, K., Lee, G. M., & Guo, Y. (2019). Gdpr-compliant personal data management: A blockchain-based solution. *IEEE Transactions on Information Forensics and Security*, 15, 1746-1761.

⁷⁰⁸ Liu, B., Ding, M., Zhu, T., Xiang, Y., & Zhou, W. (2019). Adversaries or allies? Privacy and deep learning in big data era. *Concurrency and Computation: Practice and Experience*, 31(19), e5102.

⁷⁰⁹ Ma, C., Li, J., Ding, M., Yang, H. H., Shu, F., Quek, T. Q., & Poor, H. V. (2020). On safeguarding privacy and security in the framework of federated learning. *IEEE network*, 34(4), 242-248.

⁷¹⁰ Konečný, J., McMahan, H. B., Ramage, D., & Richtárik, P. (2016). Federated optimization: Distributed machine learning for on-device intelligence. *arXiv preprint arXiv:1610.02527*.

aggregation, at the centralized server. The purpose of FL is to cooperatively learn a global model without sacrificing data privacy directly. To this end, training is performed on multiple local datasets stored at isolated data sources, without the need for collecting and processing the training data at a centralized data server. After training with local data, the participants are only required to share gradients for model updates. Then the server aggregates the gradients and transmits the updated (global) model back to each user, with the process being repeated until convergence. Thus, in FL the (global) ML model is obtained by aggregating multiple local models, each trained on its own private dataset, meaning that aggregation occurs at the model level and not at the data level, as is the case with conventional centralized ML frameworks.

This training paradigm endows FL with distinct privacy advantages compared to conventional approaches, since the personal data is stored and processed locally while only training-related parameters are exchanged. Furthermore, since the participant ID is not required by the aggregation process, updates can be transmitted without identifying metadata, utilizing de-identification methods like anonymization⁷¹¹, obfuscation methods like differential privacy⁷¹², cryptographic techniques like HE⁷¹³, and secure multi-party computation (SMC)⁷¹⁴. However, although the original data is not explicitly shared in the FL framework, personal information can still be extracted from local training parameters⁷¹⁵. For example, it has been recently shown that gradient sharing is not as secure as it is supposed to be^{716,717}, since, under certain conditions, adversaries can reconstruct the raw training data approximately, via the so called gradient inversion (GradInv) attacks⁷¹⁸. In addition, the existence of malevolent users may induce further security issues. Therefore, the design of FL still needs further protection of parameters as well as investigations on the tradeoffs between the privacy-security-level and the system performance¹¹.

EXISTING FL FRAMEWORKS

Typically, model training in a deep learning framework constitutes a mathematical optimization problem whose goal is to infer the set of model parameters that optimally maps feature vectors (inputs) to labels (output), based on example input-output pairs. One of the most popular algorithms to perform this optimization - and by far the most common way to optimize neural networks - is gradient descent⁷¹⁹, namely, an iterative method that minimizes an objective function $J(\theta)$ parameterized by the model's parameters $\theta \in \mathbf{R}^d$, by updating the parameters in the opposite direction of the gradient of the objective function $\nabla_{\theta} J(\theta)$ w.r.t. to the parameters. In the context of deep neural networks, this optimization process usually takes the following form:

⁷¹¹ Machanavajjhala, A., Kifer, D., Gehrke, J., & Venkitasubramaniam, M. (2007). l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1), 3-es.

⁷¹² Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006, March). Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference* (pp. 265-284). Springer, Berlin, Heidelberg.

⁷¹³ Papernot, N., Abadi, M., Erlingsson, U., Goodfellow, I., & Talwar, K. (2016). Semi-supervised knowledge transfer for deep learning from private training data. *arXiv preprint arXiv:1610.05755*.

⁷¹⁴ Yao, A. C. C. (1986, October). How to generate and exchange secrets. In *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)* (pp. 162-167). IEEE.

⁷¹⁵ Phong, L. T., Aono, Y., Hayashi, T., Wang, L., & Moriai, S. (2017, July). Privacy-preserving deep learning: Revisited and enhanced. In *International Conference on Applications and Techniques in Information Security* (pp. 100-110). Springer, Singapore.

⁷¹⁶ Zhu, L., Liu, Z., & Han, S. (2019). Deep leakage from gradients. *Advances in neural information processing systems*, 32.

⁷¹⁷ Aono, Y., Hayashi, T., Wang, L., & Moriai, S. (2017). Privacy-preserving deep learning via additively homomorphic encryption. *IEEE Transactions on Information Forensics and Security*, 13(5), 1333-1345.

⁷¹⁸ Zhang, R., Guo, S., Wang, J., Xie, X., & Tao, D. (2022). A Survey on Gradient Inversion: Attacks, Defenses and Future Directions. *arXiv preprint arXiv:2206.07284*.

⁷¹⁹ Rudner, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.

$$\theta_{k+1} = \theta_k - \eta \nabla \frac{1}{m} \sum_{i=1}^m L(f(x_i; \theta_k) - y_i),$$

where θ_k is the parameter vector at the k -th iteration, η refers to the learning rate hyper-parameter, L represents a loss function such as mean-square error (MSE) or the cross-entropy loss, $f(x_i; \theta_k)$ denotes the output of the model (at the k -th iteration) when presented with the i -th training input, y_i is the desired output for x_i , while $D = (x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, denotes training dataset (namely data samples for which the desired output is known). For example, in an image classification task, the pair (x_i, y_i) corresponds to the i -th input image and the corresponding ground-truth class, respectively. The parameters update process is iteratively carried out until either an acceptable local minimum is found, or the convergence speed has dropped below a pre-selected threshold.

Gradient descent methods can be further categorized into three broad categories, based on the amount of training data used in the gradient calculation at each step²¹, namely batch gradient descent, in which every update incorporates the gradients of the entire dataset D , stochastic gradient descent (SGD), where the update is performed based on the gradient of a single randomly selected sample, and mini-batch gradient descent in which the update is performed based on subsets of the training dataset, or, mini-batches (one mini-batch per update). The goal of mini-batch gradient descent is to gain the best of both worlds, that is to yield better efficiency than batch gradient descent while suppressing the gradient oscillation that SGD suffers from.

Nowadays the ever-growing computational capabilities and connectivity properties of edge devices, coupled with the growing privacy concerns over personal data generated at the edge by end-users, have helped pave the way for the migration of intelligence from the cloud to the edge, via the use of decentralized AI combining ML techniques with mobile edge-computing²⁰. In this context, FL has emerged as an ML paradigm that enables collaborative model training while retaining the original personal data of the end-users, thereby suppressing privacy-related risks, since only parameters are exchanged for model aggregation and updates, and not the data themselves. Moreover, contrary to standard distributed learning whereby the local training datasets are assumed independent and identically distributed (i.i.d) and balanced in size, FL is designed to cope with highly heterogeneous training datasets (i.e. both non-i.i.d. data and highly unbalanced sizes) scattered among (potentially) millions of compute nodes.

The distributed setting for the federated optimization is formulated as follows²¹:

$$f(w) = \frac{1}{n} \sum_{k=1}^K n_k F_k(w),$$

where the local empirical loss function $F_k(w)$ is defined as:

$$F_k(w) = \sum_{i \in P_k} f_i(w),$$

²⁰ Wang, X., Han, Y., Leung, V. C., Niyato, D., Yan, X., & Chen, X. (2020). Convergence of edge computing and deep learning: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 22(2), 869-904.

²¹ Truong, Nguyen, et al. "Privacy preservation in federated learning: An insightful survey from the GDPR perspective." *Computers & Security* 110 (2021): 102402.

with $f_i(\mathbf{w})$ being a loss function, while K is the number of local nodes, $\mathbf{P}_k$ is the training dataset set of node $k \in \{1, 2, \dots, K\}$, with size $n_k = |\mathbf{P}_k|$, and $\sum_{k=1}^K n_k = n$.

Typically, training is performed via a variation of the SGD called Federated SGD along with the Federated Averaging (FedAvg) algorithm⁷²² whereby each node updates the local model multiple times, while the orchestrator server only performs aggregation of the local models. This strategy has been shown to train deep network at 100 times fewer communications compared to the initially proposed Federated Stochastic Variance Reduced Gradient (FSVRG) algorithm⁷²³, that entails computing gradients based on $\mathbf{P}_k$ data on each local node k , obtaining a weighted average of the parameters from all local nodes, and updating the parameters for each node after the round.

The workflow cycle in a centralised FL framework consists of four steps (illustrated in Figure 20) as follows:

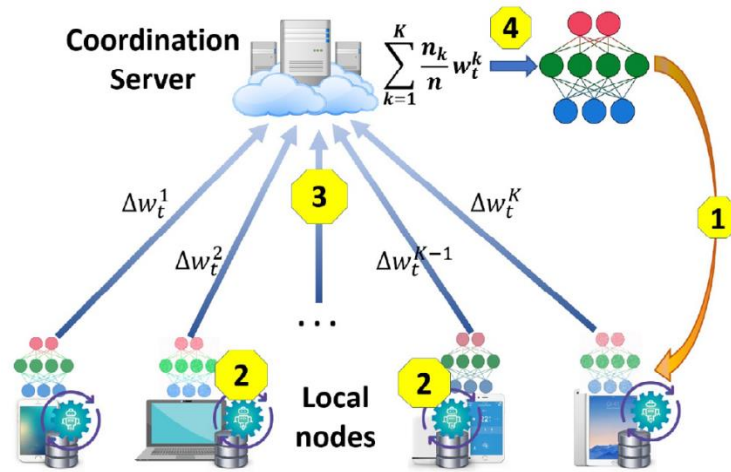


Figure 20: Centralized FL mechanism⁷²⁴

1. **Participant Selection and Global Model Dissemination:** The server selects a set of participants that satisfy requirements to be involved in the training process. It then broadcasts a global ML model (or the global model updates) to the participants for the next training round.
2. **Local Computation:** Once receiving the global ML model from the server, the participants update their current local ML model and then trains the updated model using the local dataset resided in the device. This step is operated at local nodes, and it requires end-users' devices to install an FL client program to perform training algorithms such as Federated SGD and Federated Averaging, as well as to receive the global model updates and send the local ML model parameters from/to the server.
3. **Local Models Aggregation:** The server aggregates a sufficient number of the locally trained ML models from participants in order to update the global ML model (the next step). This aggregation mechanism is required to integrate some privacy-preserving techniques such as secure aggregation, differential privacy, and advanced encryption methods to prevent the server from inspecting individual ML model parameters.

⁷²² McMahan, H. B., Moore, E., Ramage, D., & y Arcas, B. A. (2016). Federated learning of deep networks using model averaging. *arXiv preprint arXiv:1602.05629*, 2.

⁷²³ Konečný, J., McMahan, H. B., Ramage, D., & Richtárik, P. (2016). Federated optimization: Distributed machine learning for on-device intelligence. *arXiv preprint arXiv:1610.02527*.

⁷²⁴ Truong, Nguyen, et al. "Privacy preservation in federated learning: An insightful survey from the GDPR perspective." *Computers & Security* 110 (2021): 102402.

4. Global Model Update: The server performs an update on the current global ML model based on the aggregated model parameters obtained in step 3. This updated global model will be disseminated to participants in the next training round.

This 4-step cycle is repeated until the global model has reached sufficient accuracy.

COMPARATIVE STUDY OF FEDERATED LEARNING FRAMEWORKS

Table1 summarizes and highlights the main differences between FL frameworks. This study is based on recent publications^{725 726 727 728 729 730}.

The comparison criteria are as follows:

- Usability: provides information regarding the link of FL, by whom is developed, the type of license and number of stars on GitHub. Number of stars could be considered as a popularity ranking.
- Privacy: provides data privacy measures and precautions taken by FL frameworks in server and client sides such as the privacy of the model aggregation in server and the input parameters in client.
- Security: highlights the use of homographic encryption in FL frameworks.
- Data type: highlights the type of data which can be used by the FL frameworks. We focus on the popular type of data: images, tabular, or a graph.
- Mode: shows the mode of execution of a federated learning model. An FL model could run on a simulation using dockers or on real-world devices.
- NN/ML-Framework: show the type of Neural Network (NN) and Machine Learning (ML) algorithms that could be used with FL frameworks.
- Data partitioning: shows the type of data partitioning method taken into consideration in the FL frameworks. Globally, there are two types of data partitioning: horizontal and vertical. Horizontal means that the dataset share the same features space but different in sample⁷³¹. Vertical means that two datasets share the same sample ID space but differ in feature space³².
- Deployment support: shows the types of deployments. While all frameworks support the running on a single-host deployment. However, several FL frameworks provide the multi-host deployment as an option for realistic FL simulation.

Table 0.13 Comparative study of FL frameworks (N/A indicates missing information).

⁷²⁵ Liu, Xiaoyuan, et al. "Unifed: A benchmark for federated learning frameworks." arXiv preprint arXiv:2207.10308 (2022).

⁷²⁶ Ludwig, Heiko, et al. "Ibm federated learning: an enterprise framework white paper v0. 1." arXiv preprint arXiv:2007.10987 (2020).

⁷²⁷ Dimitriadis, Dimitrios, et al. "Flute: A scalable, extensible framework for high-performance federated learning simulations." arXiv preprint arXiv:2203.13789 (2022).

⁷²⁸ He, Chaoyang, et al. "Fedml: A research library and benchmark for federated machine learning." arXiv preprint arXiv:2007.13518 (2020).

⁷²⁹ Beutel, Daniel J., et al. "Flower: A friendly federated learning framework." (2022).

⁷³⁰ Rodríguez-Barroso, Nuria, et al. "Federated Learning and Differential Privacy: Software tools analysis, the Sherpa. ai FL framework and methodological guidelines for preserving data privacy." Information Fusion 64 (2020): 270-292.

⁷³¹ Yang, Qiang, et al. "Federated machine learning: Concept and applications." ACM Transactions on Intelligent Systems and Technology (TIST) 10.2 (2019): 1-19.

		TF Federa ted TFF	FedJ AX	Flowe r	FedM L	PySyft	IBM federat ed- learnin g-lib	Flute	FATE	Paddle FL
Usability										
Repository		Link ⁷³²	Link ⁷³³	Link ⁷³⁴	Link ⁷³⁵	Link ⁷³⁶	Link ⁷³⁷	Link ⁷³⁸	Link ⁷³⁹	Link ⁷⁴⁰
Developer		Google	Googl e	Adap gmbh	FedM L	OpenMi ned	IBM	Micros oft	WEBa nk	Baidu
License		Apach e-2.0	Apac he-2.0	Apac he-2.0	Apac he-2.0	Apache- 2.0	Custo me/ Not open source	MIT	Apach e-2.0	Apach e-2.0
Stars (Feb, 14 th , 2023)		2k	218	1.9k	2.4k	8.6k	385	136	4.8k	428
Privacy										
Privacy protecti on against the semi- honest server	Does not require a 3rd party aggregator (vertical)	No	N/A	No	No	N/A	N/A	No	Yes	Yes
	Aggregator does not learn model param	No	N/A	No	No	N/A	N/A	No	Yes	Yes

⁷³² <https://github.com/tensorflow/federated>

⁷³³ <https://github.com/google/fedjax>

⁷³⁴ <https://github.com/adap/flower>

⁷³⁵ <https://github.com/FedML-AI/FedML>

⁷³⁶ <https://github.com/openmined/pysyft>

⁷³⁷ <https://github.com/IBM/federated-learning-lib>

⁷³⁸ <https://github.com/microsoft/msrflute>

⁷³⁹ <https://github.com/FederatedAI/FATE>

⁷⁴⁰ <https://github.com/PaddlePaddle/PaddleFL>

	(arbitrar scenario)									
	Aggregator does not learn individual model gradient (secagg)	Yes	N/A	No	Yes	N/A	N/A	No	No	Yes
Privacy protection against semi-honest peer clients	Clients do not learn anything about the model param (vertical)	N/A	N/A	N/A	No	N/A	N/A	N/A	Yes	Yes
	Clients do not learn gradients from other clients (vertical)	N/A	N/A	N/A	Yes	N/A	N/A	N/A	Yes	Yes
Privacy protection in the final model	Support training with central DP (dpsgd/gradient edits)	Yes	N/A	Yes	No	N/A	N/A	Yes	No	Yes
Security										
	Homomorphic Encryption	N/A	No	No	N/A	Yes	Yes	No	Yes	No
	SEAL (Microsoft)	N/A	No	No	N/A	Yes	Yes	No	N/A	No
Data type										
	Graph data	No	No	N/A	Yes	No	No	No	No	No

Tabular data	No	No	Yes	Yes	No	Yes	No	No	No
Image data	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mode									
Simulation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Real world devices	No	No	Yes	Yes	Yes	N/A	Yes	No	No
NN/ML-Framework									
TensorFlow	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	N/A
Keras	Yes	No	Yes	No	Yes	Yes	No	No	N/A
PyTorch	No	No	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Clustering model	No	Yes	No	No	No	Yes	No	No	No
ML library (scikit-learn)	No	No	Yes	No	No	Yes	No	No	No
Data partitioning									
Horizontal	Yes	Yes	Yes	Yes	Yes	N/A	Yes	Yes	Yes
Vertical	No	No	No	Yes	Yes	N/A	No	Yes	No
Deployment support									
Single-host simulation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Multi-host deployment (less than 16 hosts)	No	N/A	Yes	N/A	N/A	N/A	Yes	Yes	Yes

Multi-host deployment (more than 100 hosts)	No	N/A	Yes	N/A	N/A	N/A	Yes	No	Yes

DISCUSSION

FL TOOLS

The popularity of a FL frameworks such as Flower, FedML, and PySyst is due to the availability of documentations, a large community, and easy of development and simulated/emulated settings. Most of frameworks are in Apache-2.0 license offering certain flexibility. In terms of security, especially, in homomeric encryption. There are only PySyft and FATE which offer an option to work on encrypted data. PySyft works on TenSEAL⁷⁴¹ which is a library for manipulating HE operations on tensors. TenSEAL is an extension of the SEAL Microsoft library for tensor operations, by offering HE capabilities to PySyft. FATE implements secure computation protocols based on HE and multi-party computation⁷⁴². IBM federated-learning-lib also implements and works on HE data. However, the technical documentation is not well developed compared with PySyft and FATE.

All FL tools utilize image data and run deep learning models and on simulated or emulated platform. However, some FL frameworks work on tabular data, while one of them (FedML) on graph data. Moreover, only four FL frameworks can run on real devices. Finally, some frameworks can be used and combined with classical ML algorithms from the scikit-learn python library.

FAIRNESS

Fairness has recently received increasing attention⁷⁴³. Several works studied fairness notions, but the group fairness⁷⁴⁴ is the most studied one; basically, it requires the classifier to treat different groups similarly, where groups are defined with respect to sensitive attributes such as gender and race. To this relation, demographic parity is the most commonly used group fairness notion and requires that different groups are equally likely to receive desirable outcomes. The current general machine learning models foresee a number of fairness measurements and methods, as they are deployed in crucial decision-making applications such as healthcare and recruitment. In this context, FL has been viewed as a promising solution for collaboratively training ML models among multiple parties while maintaining the privacy of their local data⁷⁴⁵. However, FL frameworks posed new challenges related to collaborative fairness, since the contributions from different participants to the learning process can be substantial different, in terms of data volume and data types. Currently, most FL systems are short of collaborative fairness, as they typically allow every participant to have access to the same global model that is trained with data from all participants⁷⁴⁶. Several recent studies focused on how to train fair classifier using FL: a client-based fairness has

⁷⁴¹ <https://github.com/OpenMined/TenSEAL>

⁷⁴² <https://fate.readthedocs.io/en/latest/>

⁷⁴³ Yuchen Zeng, Hongxu Chen, Kangwook Lee University of Wisconsin-Madison *Improving Fairness Via Federated Learning* 7 Dec 2022

⁷⁴⁴ Dwork, C.; Hardt, M.; Pitassi, T.; Reingold, O.; and Zemel, R. 2012. *Fairness through awareness*. In Proceedings of the 3rd innovations in theoretical computer science conference, 214–226.

⁷⁴⁵ FairFed: Enabling Group Fairness in Federated Learning Yahya H. Ezzeldin*, 1 Shen Yan*, 1 Chaoyang He, 1 Emilio Ferrara, 1 Salman Avestimehr 1 1 University of Southern California

⁷⁴⁶ Zirui Zhou, Lingyang Chu, Changxin Liu, Lanjun Wang, Jian Pei, and Yong Zhang. 2021. Towards Fair Federated Learning. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '21), August 14–18, 2021, Virtual Event, Singapore. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3447548.3470814>

been proposed by Li⁷⁴⁷ which aims to equalize model performance by leveraging on a federated setting named client parity across different clients; a collaborative fairness by Lyu⁷⁴⁸ which aims to reward a highly-contributing participant with a better performing local model that is given to a low-contributing participant. Another recent study aims to achieving group fairness under FL⁷⁴⁹: Ezzeldin proposed FairFed, a strategy to train models via a fairness-aware aggregation method similar to FedAvg; each client performs local debiasing on its own local dataset, thus maintaining data decentralization and avoiding exchange of any explicit information of its local data composition. Other approaches⁷⁵⁰ include heuristic approaches: the first one combines Local Fair Training with Ensemble (LFT+ Ensemble) which implies training locally fair classifiers and assembling them without violating privacy policy and the second one combines always Local Fair Training with FedAvg. In this case, each data owner runs a fair learning algorithm on its own data and locally trained models are aggregated via FedAvg.

FEDERATION MECHANISMS FOR I.I.D. AND NON-I.I.D. DATA MODELS

Without loss of generality, FL can be divided into two major categories, i.e., the horizontal and vertical FL based on the data characteristics of the participated edge devices or clients⁷⁵¹. In this study, the main focus is on the horizontal FL and existing FL mechanisms to cope with the unbalanced and non i.i.d. nature of the training datasets scattered among (potentially) millions of compute nodes.

HORIZONTAL FEDERATED LEARNING - FEDAVG ALGORITHM

The horizontal or homogeneous FL assumes that the local data of the edge devices (clients) contain the same feature space (i.e., the types of the features describing the datasets) but they have different sample space. Figure 21 illustrates an example of the horizontal FL, where two clients have datasets with the same personal features (i.e., feature space) but the datasets have derived from different persons.

⁷⁴⁷ Li, T.; Sanjabi, M.; Beirami, A.; and Smith, V. 2019. Fair Resource Allocation In Federated Learning. Published as a conference paper at ICLR 2020

⁷⁴⁸ Lyu, L.; Xu, X.; Wang, Q.; and Yu, H. 2020. Collaborative fairness in federated learning. In *Federated Learning*, 189–204. Springer

⁷⁴⁹ FairFed: Enabling Group Fairness in Federated Learning Yahya H. Ezzeldin*, 1 Shen Yan*, 1 Chaoyang He, 1 Emilio Ferrara, 1 Salman Avestimehr 1 1 University of Southern California

⁷⁵⁰ Yuchen Zeng, Hongxu Chen, Kangwook Lee University of Wisconsin-Madison *Improving Fairness Via Federated Learning* 7 Dec 2022

⁷⁵¹ Zhu, H., Xu, J., Liu, S., & Jin, Y. (2021). Federated learning on non-IID data: A survey. *Neurocomputing*, 465, 371–390. <https://doi.org/10.1016/J.NEUCOM.2021.07.098>

		Features				
Samples						
	Name	Age	Sex	Height	Weight	Label
	Person A	24	Male	178	78	1
	Person B	61	Female	165	64	0
	Person C	44	Male	182	89	1
	Person D	17	Female	159	52	0
	Person E	11	Male	137	36	1
	Person F	33	Female	171	60	0

Figure 21: A demonstration of data partition in horizontal FL. In this example the two clients contain five personal features. However, each client has data for different persons. The image was retrieved from study⁷⁵².

The FedAvg⁷⁵³ algorithm constitutes the most prominent case of horizontal FL, offering an efficient and privacy preserving distributed framework to train a machine/deep learning model (i.e., the global model) across all the involved clients without sharing any sensitive information regarding the datasets of the clients. In particular, the global model, denoted as θ and the local client's models, denoted as θ_k have the same deep learning architecture containing different parameters, since each client optimizes its local model utilizing only its corresponding local and private dataset. After the local computation, the clients broadcast the updated local models to the server. Taking into account the fact that all local models share the same architecture (e.g., number of layers), the server performs some fusion rule to aggregate the local models and generate the global model, as we can see in the following equation:

$$\theta_{global} = \sum_{k=1}^K \frac{n_k}{n} \theta^k$$

where, a weighted average is performed using all the local models of the clients and the information regarding the size of the local training dataset of each client, denoted as n_k .

⁷⁵² Zhu, H., Xu, J., Liu, S., & Jin, Y. (2021). Federated learning on non-IID data: A survey. *Neurocomputing*, 465, 371–390. <https://doi.org/10.1016/J.NEUCOM.2021.07.098>

⁷⁵³ Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H.B., Patel, S., Ramage, D., Segal, A., Seth, K., 2017. Practical secure aggregation for privacy-preserving machine learning, in: proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp. 1175–1191.

Algorithm 1 FedAvg. K is the total numbers of clients; B is the size of mini-batches, T is the total number of communication rounds, E is the total local training epochs, and η is the learning rate.

```

1: Server:
2: Initialize global model  $\theta_0$ 
3: for each communication round  $t = 1, 2, \dots, T$  do
4:   Select  $m = C \times K$  clients, where  $C \in (0, 1)$ 
5:   for each Client  $k = 1, 2, \dots, m$  in parallel do
6:     Download  $\theta_t$  to Client  $k$ 
7:     Do Client  $k$  update and receive  $\theta^k$ 
8:   end for
9:   Update global model  $\theta_t \leftarrow \sum_{k=1}^m \frac{n_k}{n} \theta^k$ 
10: end for
11:
12: Client  $k$  update:
13: Replace local model  $\theta^k \leftarrow \theta_t$ 
14: for local epoch from 1 to  $E$  do
15:   for batch  $b \in (1, B)$  do
16:      $\theta^k \leftarrow \theta^k - \eta \nabla L_k(\theta^k, b)$ 
17:   end for
18: end for
19: Return  $\theta^k$ 

```

Figure 22: The FedAvg algorithm. Note that the global model at t -th communication round is denoted by θ_t and the local model of k -th client is denoted by θ^k

FEDERATED LEARNING ON NON I.I.D. DATA

Horizontal FL approaches e.g., FedAvg framework demonstrate satisfactory results on i.i.d. datasets. As we mentioned above, the FL methodologies rely on some variant of the stochastic gradient descent. To guarantee the convergence of the stochastic gradient descent and the fact that the gradient descent is the unbiased estimate of the full gradient, the i.i.d. sampling is a crucial assumption. Nonetheless, in real-world settings and applications the local datasets of the involved clients (edge devices or nodes) are often heterogeneous, thus violating the assumption about the i.i.d. data. The non i.i.d. nature of the local data significantly deteriorates the performance accuracy of the FL methodologies, as the data distribution of the local clients is notably different from the global data distribution. Thus, the global model derived from the Local Models Aggregation step, which aggregates the received local models may diverge from the centralized true global model. Note that the centralized global model derives from the training of a deep learning model over the concatenation of all the local datasets. In literature there are numerous studies aiming to cope with challenges of the non i.i.d. data.

PERSONALIZED FEDERATED LEARNING

In this body of studies, the personalized pipeline⁷⁵⁴, instead of learning only one global model, aims to generate a “personalized” model for each local client taking into account the data heterogeneity among them. In general, there

⁷⁵⁴ V. Kulkarni, Milind Kulkarni, and A. Pant. Survey of personalization techniques for federated learning. 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), pp. 794–797, 2020.

is a plethora of works that utilize this concept, thus leading to various categories of personalization, such as local fine-tuning^{755,756,757,758} (i.e., personalization using regularization), personalization layers⁷⁵⁹ and clustering⁷⁶⁰.

LOCAL FINE-TUNING

Local fine-tuning provides personalized local models via the regularization of the local models during their training. The scope of regularization is to alleviate the divergence between the global and local models. In other words, the regularization approach prevents the local models to overfit on the local training data. Several studies, consider this type of FL by defining the following optimization problem^{761,762,763,764}

$$\sum_{k=1}^K F_k(\mathbf{w}_k) + R(\boldsymbol{\Omega}, \mathbf{w}_1, \dots, \mathbf{w}_K)$$

where R is a regularizer and $\boldsymbol{\Omega}$ is employed to relate the clients' models. The most prominent methodology is to define $\boldsymbol{\Omega}$ as the global model, thus designing a new cost function at the clients' side by adding a $\mathbf{l}_2$ norm term into the following optimization problem, i.e.,

$$\text{argmin}_{\mathbf{w}_k} = f_k(\mathbf{w}_k) + \lambda \|\mathbf{w}_k - \mathbf{w}\|_2^2$$

where $\mathbf{w}_k$ denotes the local model of the k-th client, $\mathbf{w}$ is the global model and λ stands for the regularization parameter. The above procedure provides satisfactory results on numerous datasets and applications under highly non i.i.d. scenarios.

PERSONALIZATION LAYER

⁷⁵⁵ Fallah, A., Mokhtari, A., Ozdaglar, A.E., 2020. Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach, in: Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual.

⁷⁵⁶ Jiang, Y., Konečný, J., Rush, K., Kannan, S., 2019. Improving federated learning personalization via model agnostic meta learning. arXiv preprint arXiv:1909.12488.

⁷⁵⁷ Filip Hanzely, Slavomír Hanzely, Samuel Horváth, and Peter Richtárik. Lower bounds and optimal algorithms for personalized federated learning. In NeurIPS, 2020.

⁷⁵⁸ Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. Ditto: Fair and robust federated learning through. In ICML, 2021

⁷⁵⁹ Liang, P.P., Liu, T., Ziyin, L., Allen, N.B., Auerbach, R.P., Brent, D., Salakhutdinov, R., Morency, L.P., 2020. Think locally, act globally: Federated learning with local and global representations. arXiv preprint arXiv:2001.01523 .

⁷⁶⁰ Sattler, F., Müller, K.R., Samek, W., 2020. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. IEEE Transactions on Neural Networks and Learning Systems .

⁷⁶¹ Canh T Dinh, Nguyen H Tran, and Tuan Dung Nguyen. Personalized federated learning with moreau envelopes. In NeurIPS, 2020.

⁷⁶² Filip Hanzely, Slavomír Hanzely, Samuel Horváth, and Peter Richtárik. Lower bounds and optimal algorithms for personalized federated learning. In NeurIPS, 2020.

⁷⁶³ Filip Hanzely and Peter Richtárik. Federated learning of a mixture of global and local models. arXiv preprint arXiv:2002.05516, 2020.

⁷⁶⁴ Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. Ditto: Fair and robust federated learning through. In ICML, 2021

Different from the FedAvg approach, in this case each client is allowed to contain some personalized layers in the deep learning models focused on the local data distribution of the clients. Figure 23 illustrates an example of these networks, where each client model consists of personalization layers (filled blocks) and base layers. It should be highlighted that only the base layers are uploaded to the server for the global model aggregation, thus significantly reducing the communication costs, since only the base layers need to be uploaded to the server.

FedPer algorithm⁷⁶⁵ constitutes a characteristic case of these approaches. Particularly, in the FedPer model the base layers are shallow neural networks focused on capturing high-level representations of the local data and the personalization layers are deep neural networks aiming to tackle the classification problems. This methodology is able to provide state-of-the-art performance accuracy compared to the FedAvg approach on strongly non i.i.d. data⁷⁶⁶. On the contrary, Liang et al⁷⁶⁷ proposed a different methodology in which the shallow layers of the neural network constitute the personalization layers and deep layers of the neural networks are the base layers which are broadcasted to the server. Similar to FedPer algorithm this methodology outperforms the FedAvg algorithm under non i.i.d. settings.

Overall, the personalization layers methods not only provide competitive performance accuracy but also reduce the communication burden between the clients and the server, since the clients broadcast only the base layers to the server and not the whole model, thus reducing notably the sending parameters.

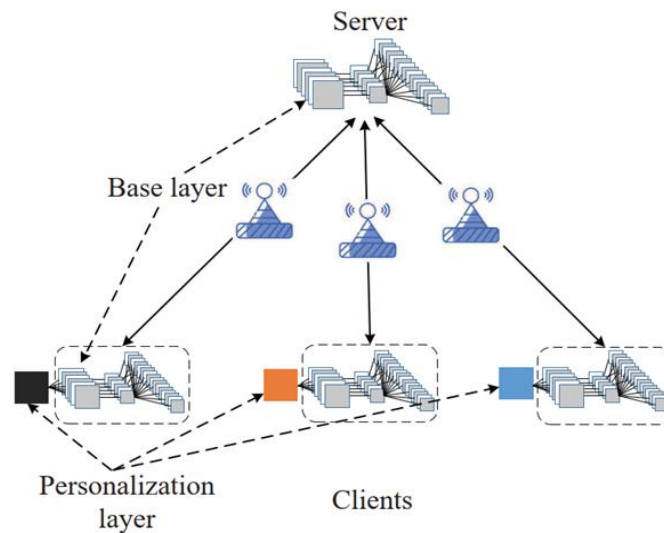


Figure 23: An illustration of horizontal federated learning with personalization layers

CLUSTERED BASED FEDERATED LEARNING – CLIENT CLUSTERING

Considering the highly Non-IID distribution of the clients' data, instead of having only one global model, a client-clustering method is proposed to develop a multi-center system. The goal is to divide the clients into different clusters based on the similarity of the local datasets. In other words, by creating multiple global models, this FL framework is able to capture the heterogeneous data distribution of the clients. The clients that belong to a specific

⁷⁶⁵ Arivazhagan, M.G., Aggarwal, V., Singh, A.K., Choudhary, S., 2019. Federated learning with personalization layers. arXiv preprint arXiv:1912.00818 .

⁷⁶⁶ Zhu, H., Xu, J., Liu, S., & Jin, Y. (2021). Federated learning on non-IID data: A survey. Neurocomputing, 465, 371–390. <https://doi.org/10.1016/J.NEUCOM.2021.07.098>

⁷⁶⁷ Liang, P.P., Liu, T., Ziyin, L., Allen, N.B., Auerbach, R.P., Brent, D., Salakhutdinov, R., Morency, L.P., 2020. Think locally, act globally: Federated learning with local and global representations. arXiv preprint arXiv:2001.01523

cluster, we assume that they follow a similar data distribution. However, since the local data are private containing sensitive information, the challenging process of the cluster-based FL approaches is to design secure data similarity evaluation frameworks to perform the clustering of the clients. In literature there are two major pipelines to securely evaluate the similarity of the nodes (clients): the first one relies on the similarity of the loss value and the second one aims to cluster the clients based on the similarity of the local models weights.

Focusing on the first similarity approach, this evaluation method is employed in several studies^{768,769,770,771}. In particular, the server generates multiple global or cluster models and broadcasts them all cluster models to the participated clients. Each client calculates the empirical loss of each cluster model based on a validation dataset, and then updates only the cluster model with the lowest loss value. After that the client uploads only this model back to the server. The major downside of this strategy is the fact that the server sends multiple models to each client at each communication round, hence heavily increasing the communication cost.

The second methodology approximates the similarity of the clients data based on the similarity of the derived local models^{772,773}. In more detail, after the server receives the local models, it calculates the similarity scores, e.g., the cosine similarity of the weights of the local models and then divides the models into clusters based on these scores. Xie et al⁷⁷⁴. defined the number of clusters and the global or cluster models before the beginning of the FL training. Additionally, this study adopted a stochastic expectation and maximization step to cluster the local clients and train the local models, simultaneously. During the expectation step, the server groups the local models into the predefined clusters. During, the maximization step, the average model of each cluster is computed, hence deriving a new global or cluster model for each cluster. After that, the derived cluster models are downloaded to the clients within the clusters for local model training.

During the inference time, each client downloads all the cluster models, and then computes the testing loss values of all cluster models, thus selecting the model with the highest performance accuracy. Sattler et al⁴⁸. proposed a different tree-based structure. At the root node resides, a global model $\mathbf{w}$ is trained with the FedAvg algorithm over all connected clients. In the next layer (the 1st split layer), the client population is split into two clusters based on the cosine similarities, and the cluster models * 0 and *1 are trained again by the FedAvg algorithm on the allocated cluster clients. This splitting continues recursively until the stop condition is satisfied. It can be seen that the cluster models near the root node are general models and those located at the bottom layers are personalized models. And different cluster models in the parameter tree can be selected for different applications. Moreover, all the tree models can also be ensemble by the aforementioned two approaches to metrics calculation. Figure 24 illustrates the above methodology.

⁷⁶⁸ Mansour, Y., Mohri, M., Ro, J., Suresh, A.T., 2020. Three approaches for personalization with applications to federated learning. arXiv preprint arXiv:2002.10619 .

⁷⁶⁹ Koppurapu, K., Lin, E., 2020a. Fedfmc: Sequential efficient federated learning on non-iid data. arXiv preprint arXiv:2006.10937

⁷⁷⁰ Ghosh, A., Hong, J., Yin, D., Ramchandran, K., 2019. Robust federated learning in a heterogeneous environment. arXiv preprint arXiv:1906.06629

⁷⁷¹ Ghosh, A., Chung, J., Yin, D., Ramchandran, K., 2020. An efficient framework for clustered federated learning. arXiv preprint arXiv:2006.04088

⁷⁷² Briggs, C., Fan, Z., Andras, P., 2020a. Federated learning with hierarchical clustering of local updates to improve training on non-iid data, in: 2020 International Joint Conference on Neural Networks (IJCNN), IEEE. pp. 1–9.

⁷⁷³ Sattler, F., Müller, K.R., Samek, W., 2020. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. IEEE Transactions on Neural Networks and Learning Systems .

⁷⁷⁴ Singh, I., Zhou, H., Yang, K., Ding, M., Lin, B., Xie, P., 2020. Differentially-private federated neural architecture search. arXiv preprint arXiv:2006.10559

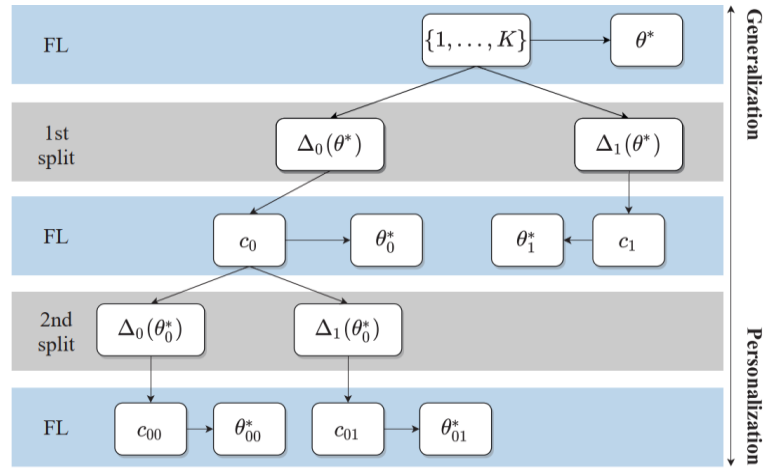


Figure 24: The parameter tree created by cluster federated learning

Overall, performing clustering on clients with heterogeneous data distributions constitutes a sensible and effective approach to cope with the non i.i.d. data problem, since the aggregation of all the local models into a single model might provoke negative knowledge transfer and deteriorate the performance of the global model. Additionally, the incorporation of multiple global models provides a flexibility on the federated learning systems, since different global models can be utilized and ensembled for various applications and settings. Nonetheless, the cluster-based FL requires considerable communication resources during the training and inference procedures.

SWARM LEARNING

Swarm Learning, developed by Hewlett Packard Enterprise⁷⁷⁵, is a decentralized, privacy-preserving framework for performing machine learning model training at the data source. This solution exploits the computational power available at or near the distributed data sources, thus it allows the data staying local, whereas only the acquired learnings are shared. This results in improved models with less bias, while at the same time solving for data privacy, data ownership, and efficiency concerns. Swarm Learning leverages a permissioned blockchain to securely onboard members, and dynamically elect the leader providing resilience and security.

Swarm Learning is in line with the key principles subsumed by TRUSTEE, in that it enables decoupling of data ownership from data access, authorization of data owners to any incoming service request, contractual ruling of the service, and assurance that raw data never leave their source place. The presence of blockchain data contracts, in turn, increased control of data owners on access and usage of their data, so improving the compliance with privacy and security regulations.

Like FL, Swarm Learning represents an alternative approach to centralized Machine Learning. It does not require to send and aggregate training data coming from peripheral data sources into a centralized location to perform model training and optimization. The differentiation respect to FL stays in Swarm Learning not even sending model parameters to a centralized processing node (central custodian). Swarm Learning is a fully decentralized solution, exploiting the computational power at – or close to – the actual data sources. In a Swarm Learning system, both the core operations, i.e. model training and subsequent inferencing through application of the trained model, occur at the edge location, where the data are produced, and where data driven decisions and actions must actually be enforced. It is a fully peer-to-peer solution, where raw data stay still at their originating location, and only the

⁷⁷⁵ <https://www.hpe.com/us/en/hpe-swarm-learning.html>

learned insights are shared among the collaborating nodes, with related advantages in terms of data security and privacy too.

Experiments have shown a >1000X reduction of exchanged information compared to raw data transfer, bringing equivalent reduction of cost and processing delays. Swarm Learning reduces the latency between data creation and availability of actionable insights, since it allows to trigger model retraining as long as new data become available at one of the sources, without having to wait full blocks of data transfer, consolidation and mining. This solution also avoids the costs of infrastructure needed to deploy a centralized server. Finally, a fully distributed ML gets the system completely rid of any potential single point of failure, so improving its fault-tolerance grade.

Swarm Learning was ultimately inspired by biology. It is a decentralized ML solution built on blockchain technology, designed to address also enterprise-grade environments, with the objective to harness the power of distributed data ensuring proper privacy and security levels. As said, it ensures privacy by sharing only insights derived from data rather than raw data itself. The Swarm Learning framework ensures that only legitimate participants can join the decentralized learning network, and a smart contract regulates the contributions and rewards from/to participants. The smart contract is modelled in an innovative shape, accompanied by a monetization framework to facilitate multi-party collaboration in terms of business relationship whenever this latter is involved.

Swarm Learning enables a set of nodes, each one owning some local training data, to cooperatively train a common ML model with zero sharing of data. This is achieved by single nodes sharing some coefficients (weights) obtained by the local training performed on the model by their local data. So, peer nodes can keep full privacy of their own data. Parameters shared by the peer nodes are then merged to come up with a global model. The merge however does not happen at a static central coordinator or parameter server. A temporary leader is dynamically selected for this action among the peers themselves, also offering a better fault-tolerance. This mechanism makes the full “collective” intelligence of the network available to each peer node, without requiring the data to ever move from their source. The model training leverages Distributed ML technology (usually named *data parallel paradigm* in ML). Blockchain technology solves the inter-domain decentralized control, scalability and fault-tolerance requirements, plus offering a tamperproof cryptocurrency framework enabling the peer nodes to monetize their contributions.

An extensive study published in January 2022⁷⁷⁶ made an accurate assessment of Swarm Learning performance in terms of prediction accuracy. The researchers reached the conclusion that Swarm Learning accuracy is similar to the one achievable by centralized learning, and shows even more fairness in case of datasets biased over not highly relevant features.

Swarm nodes can generally be deployed on any kind of infrastructure, more or less powerful. Each Swarm node includes a SPIRE⁷⁷⁷ agent Workload Attestor plugin, communicating with SPIRE server nodes to attest their identity, and acquire a SPIFFE Verifiable Identity Document (SVID), so solving reciprocal authentication issues.

FUNCTIONAL FLOW

A Swarm Learning workflow is made up by three operational phases:

1. Initialization & onboarding
2. Installation/configuration
3. Integration and training

⁷⁷⁶ <https://arxiv.org/pdf/2201.05286.pdf>

⁷⁷⁷ <https://github.com/spiffe/spire>

Onboarding is the (offline) phase when multiple entities actually join a Swarm network and define the operational and legal rules for their common decentralized network. The rules encompass data/data parameter sharing agreements, provisions to ensure the visibility of nodes across administrative domain boundaries, consensus on expected outcome of the model training. Onboarding also encompasses definition of values for configurable parameters, individuation of peer-discovery nodes supporting the boot-up, and specification of inter-node synchronization frequency. As last step, the agreement on the common model to train and the rewarding system (if present) are finalized.

At onboarding completed, all the peers download and install the Swarm platform software on their nodes. The software comes along with the previously agreed Swarm network configuration. Once installed, the platform boots up starting its connection to the Swarm network: this is ultimately a blockchain overlaid on the underlying network connection among nodes. The boot-up is sequential, starting from the subset of nodes appointed as peer-discovery nodes, then going on with the others.

Swarm Learning exposes a simple API, enabling integration with multiple frameworks, embedded within the code base. Through the API, a stand-alone node is easily turned into a Swarm learning peer. Then, the model training phase comes in, divided in five steps (Figure 25):

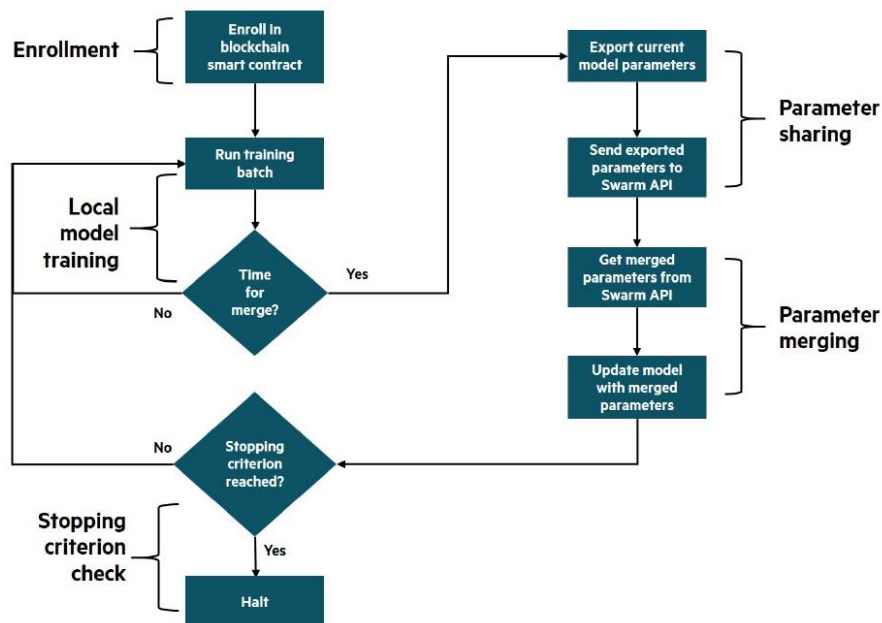


Figure 25: Swarm Learning training steps⁷⁶

1. *Enrollment* (registration): every node subscribes, once for good, to the acting smart contract. Afterwards, the node records in the contract its key attributes, like the URI accessible by other nodes to download its trained parameters.
2. *Local model training*: nodes go on to train the model's local copy, in a multi-step iterating manner, where each round is named *epoch*. Inside an epoch, the node trains its local model by one or more data batches for a given number of iterations. Once the target number of iterations is done, the node packs the obtained values of parameters into a file and exports them uploading to a shared file system, where other nodes can read it. Afterwards, the node signals to its peers its readiness for the parameter sharing phase.
3. *Parameter sharing*: this phase is triggered by a minimum number of nodes signaling to be ready; the threshold number of nodes is defined in the initialization phase. First step is the election of an *epoch leader*, performed by a predetermined election algorithm. The epoch leader takes over the merging of parameter values uploaded by nodes as result of their local training. This is a fast step, occurring at the end of an

epoch. In this phase, the system is typically shaped as a star topology, nevertheless other topologies (like k-way merge, where a set of nodes carries out the merge) can easily be configured. Exchanged parameters can be, and usually are, encrypted before being transferred.

4. *Parameter merging*: the framework can support different merge algorithms, e.g., mean, weighted mean, median, etc. By means of the chosen algorithm, the epoch leader puts together the values uploaded by all the nodes, packages a new file containing the merged values, and communicates to the other nodes the availability of this file. Every node then downloads the merged file and updates its local model with the merged parameter values.
5. *Verification of stopping criterion*: the nodes re-evaluate the model applying the updated parameter set, again using their local datasets to compute a number of possible validation metrics. The values out of this step are again shared with other nodes, by means of the state variable in the smart contract. Once a node is done with this step, it signals to the network its completion. In the meanwhile, the leader polls the completion signals from all nodes, and when he verifies that all the nodes are done, it merges the local validation metric values to compute a set of global metric values. This makes the synchronization completed for good, and is also the point where, if a monetization has been agreed, the rewards for each participant are computed and allotted.

At this point, the system status is matched with the stopping criterion. If the matching is met, the Swarm Learning process can stop here. If there is no matching yet, steps 2-5 are replicated until the criterion gets eventually fulfilled.

ARCHITECTURE

The architecture of Swarm Learning is four-layered:

- API layer
- Control layer
- Data layer
- Monetization layer

These components are modular, so they can be re-implemented with different technologies without hindering the overall framework functionality. The framework can run on differently powered hardware and heterogeneous infrastructures. It can be deployed on one or more datacenters and is fault tolerant respect to the warm joining/disjoining of nodes, which does not compromise the whole network's operation. Figure 26 depicts the Swarm Learning architecture.

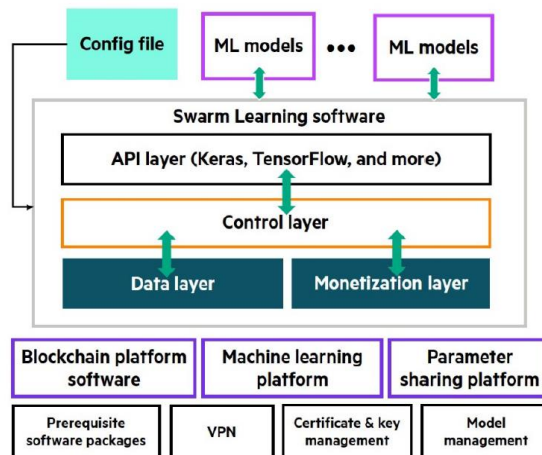


Figure 26: Swarm Learning architecture

Swarm learning implementation format is an API library, supporting various frameworks among which TensorFlow, Keras and others. The interface provided by the API resembles the one offered by the training APIs of mainstream frameworks. API calls create the hook points for Swarm Learning, enabling the node to exchange parameters at every epoch's end. With few simple code updates, we can obtain a cohort-learning configuration, hiding any complexity of control and data flow.

The control layer oversees ensuring a consistent state to the whole Swarm network. It is based on blockchain technology. This layer ensures atomicity of operations and state transitions, encapsulating state and operations in a smart contract. The state information includes current epoch, Swarm members with IP address and ports, URIs of parameter files. The operations include the logic of leader election, fault-tolerance mechanisms, inter-node signaling.

The data layer is responsible for reliable and secure parameter sharing. Even this layer is pluggable, supports different file-sharing mechanisms (HTTPS/TLS, IPFS, and others). The data layer is controlled by the control layer, which also stores all its key information.

Monetization layer performs data usage and contributions metering in the model training, to compute monetary rewards for the participants, dished out at the end of the training phase. It is based on blockchain tamperproof and self-verifying smart contracts, keeping track of contributions, and a built-in cryptocurrency framework, to automate reward transfers.

EXEMPLARY APPLICATION AREAS

One particularly significant case story about Swarm Learning in the healthcare/life science research domain is given by its usage at DZNE (German Center for Neurodegenerative Diseases) in supporting clinical research on brain pathologies (more information [here](#)). In general, healthcare is one of the domains offering the most interesting opportunities, given that medicine is intrinsically a decentralized domain. There is a frequent situation of “bias” in hospitals, meaning that each hospital detains a dataset biased toward the most common illnesses present in its coverage area, and has less information to use on less occurring diseases. Sharing data among different hospitals can allow each one to fill its own gaps with the data coming from other hospitals. Research on precision medicine and fast detection of severe illnesses, for instance, is often conducted in collaboration among worldwide research institutions, which implies strong regulatory and compliance issue since it involves highly sensitive data. Last but not least, patient records data are substantially not standardized, so sharing learnings in a disjoint manner poses problems of data semantic interpretation. Swarm learning can clear out these problems by eliminating raw data movements, and not by chance there is already a very extensive set of use cases where it has been successfully experimented.

The paper at⁷⁷⁸ presents in deep detail four use cases of heterogeneous diseases (COVID-19, tuberculosis, leukemia and lung pathologies) where Swarm Learning was applied to develop disease classifiers based on blood transcriptomes, with results outperforming what would be achievable by single node processing.

Another case study was on skin lesion classification by dermoscopic images⁷⁷⁹, which performed a comparative analysis of fairness among different training approaches. Fairness has been raised as a possible critical point in FL systems⁷⁸⁰, since they are considered vulnerable to biases in the model, due to the model assigning more relevance to dominant subgroups, like users with sensitive attributes. This can lead to serious consequences once the obtained

⁷⁷⁸<https://www.nature.com/articles/s41586-021-03583-3>

⁷⁷⁹<https://arxiv.org/pdf/2109.12176.pdf>

⁷⁸⁰ Li, T., Sanjabi, M., Beirami, A., & Smith, V. (2019). Fair resource allocation in federated learning. *arXiv preprint arXiv:1905.10497*.

results are used to take clinical decisions. The study concluded that Swarm learning keeps up with centralized learning and outperforms single-node training (Figure 27). Another case is about histology image analysis, described in⁷⁸¹.

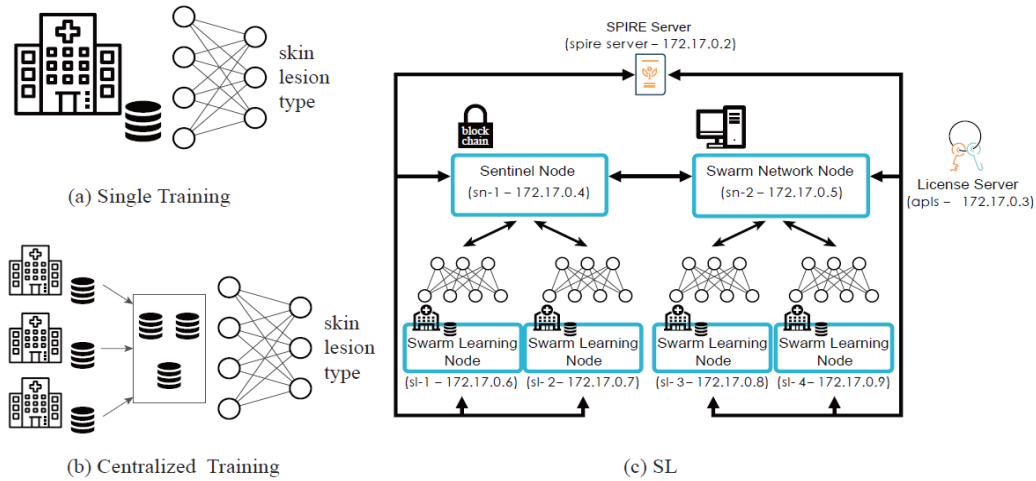


Figure 27: Different training strategies⁷⁸²

Urban mobility is another domain rich of opportunities. Combining rich, heterogeneous datasets (e.g., commuting paths, daily routine paths, road information, weather information, events, etc.) we can achieve more accurate traffic predictions, and optimized public transportation planning. Interconnected cars are another very promising area, where the need to quickly process events and get actionable learnings is paramount.

Space scenarios are also highly interesting, being them cases where the latency among data sources can lift up to huge values (tens of minutes between Earth and Mars).

A community version of the Swarm Learning software is available at Github⁷⁸³.

DATA LEAKAGE IN FEDERATED LEARNING

PRIVACY VS SECURITY

Here, we highlight the differences between terms “privacy” and “security”¹¹:

- *Security issues*: unauthorized/malicious access, change or denial to data. Such attacks are usually launched by hackers with expert knowledge of the target system or network. Hence, the fundamental three goals of security are confidentiality, integrity, and availability.
- *Privacy issues*: unintentional disclosure of personal information, usually from open-access data. For example, from a side-by-side comparison of a vote registration dataset and an anonymous set of healthcare sensor records (e.g., no individual’s name and ID), an attacker may be able to identify certain individuals and learn about their health conditions. This is because some quasi-identifiers such as gender, birth date, and zip code are the same in both datasets.

⁷⁸¹ <https://www.nature.com/articles/s41591-022-01768-5>

⁷⁸² <https://arxiv.org/pdf/2109.12176.pdf>

⁷⁸³ <https://github.com/HewlettPackard/swarm-learning/releases/tag/v1.2.0>

Privacy attacks only require common sense and involve no hacking activities. The fundamental reason of privacy issues is that a seemingly harmless open dataset may contain clues to an individual's private information in real life. Hence, alternative goals such as anonymity, unlinkability, and unobservability have been proposed for privacy protection.

PRIVACY ISSUES OF FEDERATED LEARNING

Based on the works in^{784,77,78}, we highlight three types of privacy issues of FL:

- a. *Unintentional Data Leakage and Model Reconstruction Attacks*. This issue is related to the model. In fact, the data could be inferred by only considering model weights where an honest but curious client may eavesdrop communications between clients and the server by reconstructing the client model⁷⁸⁵. In ⁷⁸⁶ ⁷⁹ demonstrated the possibility of reconstructing local data by inverting the model gradients sent by the clients to the server. The authors in ⁷⁸⁷ proposed and evaluated several inference attacks. These attacks enable a malicious user to infer the presence of exact data points in other clients' training data, as well as properties that characterize subsets of the training data.
- b. *Malicious Server*. In a centralized FL architecture, the central server is the most critical parts of the architecture. The server could be a malicious or accessed by unauthorized users. Then, all clients' models communications with this server can be intercepted and studied for reconstructing a part of the original client data ⁷⁸⁸.
- c. *Passive and Active Inference Attack*. The passive attack aims to analyze the user inputs at the server level. The active attack works on sending the global update to only an isolated client. Experiments in ⁷⁸⁹ demonstrated that with only some of the computed gradients of a neural network, training data could be inferred. The authors were able to reconstruct nearly the whole original data by inference. In ⁷⁹⁰ the authors experimented passive inference attack called a *passive membership inference attack*. This attack allows to the adversary inferring whether the client owns a specific data instance. It allows to determine if data have already been seen by the model. This type of attack exploits the stochastic gradient descent algorithm to extract information about training sets.

SECURITY ISSUES OF FEDERATED LEARNING

POISONING ATTACKS ON FL

⁷⁸⁴ Gosselin, Rémi, et al. "Privacy and Security in Federated Learning: A Survey." *Applied Sciences* 12.19 (2022): 9901.

⁷⁸⁵ Nasr, Milad, Reza Shokri, and Amir Houmansadr. "Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning." *2019 IEEE symposium on security and privacy (SP)*. IEEE, 2019.

⁷⁸⁶ Geiping, Jonas, et al. "Inverting gradients-how easy is it to break privacy in federated learning?." *Advances in Neural Information Processing Systems* 33 (2020): 16937-16947.

⁷⁸⁷ Melis, Luca, et al. "Exploiting unintended feature leakage in collaborative learning." *2019 IEEE symposium on security and privacy (SP)*. IEEE, 2019.

⁷⁸⁸ Song, Mengkai, et al. "Analyzing user-level privacy attack against federated learning." *IEEE Journal on Selected Areas in Communications* 38.10 (2020): 2430-2444.

⁷⁸⁹ Zhao, Bo, Konda Reddy Mopuri, and Hakan Bilen. "idlg: Improved deep leakage from gradients." *arXiv preprint arXiv:2001.02610* (2020).

⁷⁹⁰ Zari, Oualid, Chuan Xu, and Giovanni Neglia. "Efficient passive membership inference attack in federated learning." *arXiv preprint arXiv:2111.00430* (2021).

One of the privacy-preserving objectives of centralized FL is that a coordination server is unable to inspect the data or administer the training process at a local node. This, however, prohibits the transparency of the training process; thus, imposes a new vulnerability of a new type of attack called model poisoning^{791,792,793,794,795}. Generally, model poisoning attacks aim at manipulating the training process by feeding poisoned local model updates to a coordination server. This type of attack is different from data poisoning^{796,797}, which is less effective in FL settings^{83,798} because the original training data is never shared with a server. Thus, this section is mainly dedicated to analyzing the model poisoning attacks in FL. Generally, model poisoning is conducted at the client-side wherein an adversary controls a fraction of participants for a common adversarial goal, either (i) corrupting the global model so that it converges to a sub-optimal which is an incompetent, ineffective one (i.e., random attack)^{84,85}, or (ii) replace it to a targeted model (i.e., replacement attack)^{83,90}. Compromised participants analyze the targeted global model; the poisoned model is then trained on backdoor data samples using dedicated techniques such as constrain-and-scale accordingly and feed the parameters to a coordination server as other honest participants. The objective of this attack is that the global model is replaced by a joint model consisting of both the original task and the injected backdoor sub-task while retaining high accuracy on the two. The backdoor training at the adversary can be empowered by modifying minimization strategies such as constrain-and-scale, which optimizes both gradients of the loss and the backdoor objective⁹⁰. A parameter estimation mechanism is then used for generating parameters submitted to the coordination server for honest participants' updates. As secure aggregation is used for preventing the server from inspecting individual models, this poisoning model is unable to detect^{83,90}.

INFERENCE ATTACKS ON FL

As aforementioned, a trained ML model contains unintended features that can be utilized to extract personal information. Thus, local ML model parameters from a federated optimization algorithm can be exploited by an adversary to infer personal information, particularly when combining with related information such as model data structure and meta-data. This information can be either original training data samples (i.e., reconstruction attack)^{19,90,77,78} or membership tracing (i.e., to check if a given data point belongs to a training dataset)^{79,799}.

Attackers might carry out model inversion (MI) attack to extract sensitive information contained in training data samples, for instance, by reconstructing representatives of classes which characterizing features in classification

⁷⁹¹ Bhagoji, A. N., Chakraborty, S., Mittal, P., & Calo, S. (2019, May). Analyzing federated learning through an adversarial lens. In *International Conference on Machine Learning* (pp. 634-643). PMLR.

⁷⁹² Blanchard, P., El Mhamdi, E. M., Guerraoui, R., & Stainer, J. (2017). Machine learning with adversaries: Byzantine tolerant gradient descent. *Advances in neural information processing systems*, 30.

⁷⁹³ Chen, L., Wang, H., Charles, Z., & Papailiopoulos, D. (2018, July). Draco: Byzantine-resilient distributed training via redundant gradients. In *International Conference on Machine Learning* (pp. 903-912). PMLR.

⁷⁹⁴ Fung, B. C., Wang, K., Chen, R., & Yu, P. S. (2010). Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys (Csur)*, 42(4), 1-53.

⁷⁹⁵ Guerraoui, R., & Rouault, S. (2018, July). The hidden vulnerability of distributed learning in byzantium. In *International Conference on Machine Learning* (pp. 3521-3530). PMLR.

⁷⁹⁶ Jagielski, M., Oprea, A., Biggio, B., Liu, C., Nita-Rotaru, C., & Li, B. (2018, May). Manipulating machine learning: Poisoning attacks and countermeasures for regression learning. In *2018 IEEE symposium on security and privacy (SP)* (pp. 19-35). IEEE.

⁷⁹⁷ Koh, P. W., & Liang, P. (2017, July). Understanding black-box predictions via influence functions. In *International conference on machine learning* (pp. 1885-1894). PMLR.

⁷⁹⁸ Bagdasaryan, Eugene, et al. "How to backdoor federated learning." *International Conference on Artificial Intelligence and Statistics*. PMLR, 2020.

⁷⁹⁹ Bonawitz, Keith, et al. "Towards federated learning at scale: System design." *Proceedings of machine learning and systems* 1 (2019): 374-388.

ML models⁸⁰⁰. MI attacks do not require the attacker to actively participate in the training process (i.e., black-box or passive attacks). For example, it is possible to recover images from a facial recognition model for a particular person (i.e., all class members depict this person) using MI by deriving a correct weighted probability estimation for the target feature vectors⁷⁸. In this scenario, the experiment results show that this MI attack can reconstruct images that are visually similar to the victim's photos⁹².

In an FL framework, attackers are not only able to observe the trained model parameters but also participate in the training process to inspect the changes in the updated global models in some consecutive training rounds (i.e., white-box or active attacks), which will intensify the.

It is shown that MI attacks based on class representation are more challenging than reconstructing from gradients for classification models⁷⁸. In this regard, numerous reconstruction attacks were proposed based on Generative Adversarial Networks (GANs)⁸⁰¹ to synthesize fake samples which have the same statistics (e.g., distribution) to those in the training set without having access to the original data. For instance, it has been shown that by utilizing GANs, an insider can infer information from a victim just by analyzing the shared model parameters in some consecutive training rounds⁸⁰². This attack can be accomplished at the client-side without interfering the whole FL procedure, even when the local model parameters are obfuscated using the DP technique. A malicious coordination server can also recover partial personal data by inspecting the proportionality between locally trained model parameters sent to the server and the original data^{19,803}. However, it is not necessary to fully reconstruct the trained data; instead, inferring attributes or membership of the original trained data from local model parameters can also induce serious privacy leakage^{76,78,804} (e.g., an attacker can figure out whether a specific data sample (of a patient) is used to train a model of a disease). This is the baseline for the membership attack. Authors have also investigated membership attacks in FL and demonstrated the capability of these attacks in both passive and active approaches^{76,78}. For instance, the gender of a victim can be inferred with very high accuracy of 90% when conducting this attack in a binary gender classifier on the Face-Scrub dataset⁸⁰⁵. Other features, which are uncorrelated with the main task, can also be inferred such as race and facial appearance (e.g., whether a face photo is wearing glasses)⁷⁸, while proposed an active attack approach called gradient ascent by exploiting the privacy vulnerabilities of SGD optimization algorithms has also been proposed⁷⁶. This attack based on the correlation between the local gradients of the loss and the direction and the amount of changes of model parameters when minimizing the loss to fit a model to train data samples in the SGD algorithms. This active membership attack was conducted on the CIFAR100 dataset⁸⁰⁶ showing a high accuracy of 74% compared to only 50% in passive attack⁷⁶.

BACKDOOR ATTACKS.

⁸⁰⁰ Fredrikson, M., Jha, S., & Ristenpart, T. (2015, October). Model inversion attacks that exploit confidence information and basic countermeasures. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security* (pp. 1322-1333).

⁸⁰¹ Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144.

⁸⁰² Hitaj, Briland, Giuseppe Ateniese, and Fernando Perez-Cruz. "Deep models under the GAN: information leakage from collaborative deep learning." *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*. 2017.

⁸⁰³ Wang, Zhibo, et al. "Beyond inferring class representatives: User-level privacy leakage from federated learning." *IEEE INFOCOM 2019-IEEE conference on computer communications*. IEEE, 2019.

⁸⁰⁴ Ganju, Karan, et al. "Property inference attacks on fully connected neural networks using permutation invariant representations." *Proceedings of the 2018 ACM SIGSAC conference on computer and communications security*. 2018.

⁸⁰⁵ Ng, Hong-Wei, and Stefan Winkler. "A data-driven approach to cleaning large face datasets." *2014 IEEE international conference on image processing (ICIP)*. IEEE, 2014.

⁸⁰⁶ Krizhevsky, Alex, and Geoffrey Hinton. "Learning multiple layers of features from tiny images." (2009): 7.

A single or multiple malicious participants can introduce back-door functionality to perform a Data Poisoning into the joint model, e.g., modify the label to a misclassification. The authors in⁹⁰ show the possibility to create a backdoor by poisoning the models. It introduces a specific bias focused on certain labels. A classic example of a backdoor in FL is the image detection models⁸⁰⁷, where the image classifier is distorted to assign a specific label chosen by the attacker.

MITIGATION MECHANISMS WITH EMPHASIS IN HE-ENABLED FL

In recent years, the privacy protection of sensitive data has become a crucial research area in deep learning since the processing of sensitive data raises concerns about the privacy and security. FL as a private and distributed framework enables multiple parties to collaborate on a machine learning model without sharing their data. However, the FL methodology requires the participated devices to broadcast their trained models, which may disclose some sensitive information about the local datasets⁸⁰⁸. To ensure privacy in the context of FL, it is important to prevent adversaries from inferring sensitive data from model updates transmitted during the training process. Broadcasting these updates in plaintext can put the privacy of the data at risk. One efficient approach to address this issue is to use HE⁸⁰⁹, which enables secure computation on encrypted data.

Homomorphic encryption and Federated learning

In literature several studies have explored the use of HE in FL and have demonstrated its effectiveness in improving data privacy and security. Hao et al⁸¹⁰. proposed a privacy-preserving FL methodology utilizing differential privacy and symmetric additive HE. However, the method suffered from a loss of accuracy due to the introduction of differential privacy noise. Zhang et al⁸¹¹. developed a multipart deep learning approach that used asynchronous optimization, lightweight HE, and threshold secret sharing. This approach achieved strong privacy guarantees, desirable model utility, and a balance between model utility and training efficiency. Fang et al⁸¹². also proposed a privacy-preserving and communication-efficient approach for FL in the IoT that combines secret sharing with lightweight HE and is resistant to various collusion scenarios. While Wang et al⁸¹³. proposed a privacy-friendly data collection scheme for Deep Learning using FL, there were still potential privacy threats in FL. To address these issues, Hao et al⁸¹⁴. presented an efficient and privacy-friendly FL scheme for industrial artificial intelligence using an improved homomorphic BGV encryption scheme and differential privacy. Their approach was non-interactive in any aggregation, providing a high level of privacy for both local and shared parameters.

⁸⁰⁷ Zhang, Jiale, et al. "PoisonGAN: Generative poisoning attacks against federated learning in edge computing systems." *IEEE Internet of Things Journal* 8.5 (2020): 3310-3322.

⁸⁰⁸ A. Bhowmick, J. Duchi, J. Freudiger, G. Kapoor, R. Rogers, Protection against reconstruction and its applications in private federated learning, 2018, arXiv preprint arXiv:1812.00984.

⁸⁰⁹ Aono, Y., Hayashi, T., Wang, L., & Moriai, S. (2017). Privacy-preserving deep learning via additively homomorphic encryption. *IEEE Transactions on Information Forensics and Security*, 13(5), 1333-1345.

⁸¹⁰ Hao, M., Li, H., Xu, G., Liu, S. & Yang, H. Towards efficient and privacy-preserving federated deep learning, 1–6 (IEEE, 2019).

⁸¹¹ Zhang, X., Ji, S., Wang, H. & Wang, T. Private, yet practical, multipart deep learning, 1442–1452 (2017). <https://doi.org/10.1109/ICDCS.2017.215>.

⁸¹² . Fang, C., et al. (2021). Privacy-preserving and communication efficient federated learning in internet of things. *Computers & Security*, 103, 102199. <https://doi.org/10.1016/j.cose.2021.102199>

⁸¹³ Wang, T., et al. (2020). Privacy-enhanced data collection based on deep learning for internet of vehicles. *IEEE Transactions on Industrial Informatics*, 16(10), 6663–6672. <https://doi.org/10.1109/TII.2019.2962844>

⁸¹⁴ Hao, M., et al. (2020). Efficient and privacy-enhanced federated learning for industrial artificial intelligence. *IEEE Transactions on Industrial Informatics*, 16(10), 6532–6542. <https://doi.org/10.1109/TII.2019.2945367>

Additionally, Wibawa⁸¹⁵ et al presented a privacy-preserving FL algorithm for COVID-19 detection, which involved two phases: local model training at each client and encrypted model weight aggregation at the server using a BFV crypto-scheme. The proposed algorithm employed a secure multi-party computation protocol to protect the deep learning model from adversaries. The algorithm's performance was evaluated using a real-world medical dataset, demonstrating its ability to improve medical data privacy and security while maintaining model accuracy. Li et. al.⁸¹⁶ focused on preserving the identity privacy of the drivers through the use of FL in autonomous driving model training. In particular, they adopted a HE algorithm, namely, the 'Improved Dijk-Gentry-Halevi-Vaikuntanathan (DGHV) algorithm,' to enhance message privacy. The algorithm was successfully used to protect the message information in the proposed traceable identity-based scheme under the assumption of an honest-but-curious MEC server and malicious vehicles.

Finally, Zhang et. al⁸¹⁷. introduced a masking scheme for FL that utilized HE. The proposed method aimed to address the issue of data quality being overlooked in traditional model aggregation algorithms by implementing a weighted average algorithm based on data quality. This approach resulted in reduced negative impacts caused by data heterogeneity on the global model's accuracy and convergence rate. Additionally, the proposed scheme utilized HE to encrypt only a variable called data quality for each client, thus minimizing computation overhead.

A generic HE-Federated learning scheme^{818 819 820}

Set up:

1. Parties or clients share a HE private key for the decryption process, and share with the server a public key for the encryption process
2. The server and clients define the deep learning model architecture
3. Each client trains the deep learning model locally, using its private data, and produces a set of model updates. The model updates are then encrypted using the shared HE public key, so that they can be sent securely to the server.
4. Once the model updates have been encrypted, each client broadcasts its encrypted model update to the server. This can be done using a secure communication channel, such as a secure socket layer (SSL), to ensure that the updates are not intercepted or modified during transmission.
5. After receiving the encrypted model updates from the clients, the server applies a fusion/aggregation rule, such as averaging or weighted averaging, to the encrypted updates to obtain a new encrypted global model. The fusion/aggregation rule used depends on the specific application and the desired level of accuracy.
6. After that the server sends to clients the new encrypted model. The clients decrypt the global model and use it to update their local models for the next round of training. The process is repeated until convergence.

⁸¹⁵ Wibawa, Febrianti, et al. "Homomorphic encryption and federated learning based privacy-preserving cnn training: Covid-19 detection use-case." *Proceedings of the 2022 European Interdisciplinary Cybersecurity Conference*. 2022.

⁸¹⁶ Li, Yijing, et al. "Privacy-preserved federated learning for autonomous driving." *IEEE Transactions on Intelligent Transportation Systems* 23.7 (2021): 8423-8434.

⁸¹⁷ L. Zhang, J. Xu, P. Vijayakumar, P. K. Sharma and U. Ghosh, "Homomorphic Encryption-based Privacy-preserving Federated Learning in IoT-enabled Healthcare System," in *IEEE Transactions on Network Science and Engineering*, 2022, doi: 10.1109/TNSE.2022.3185327.

⁸¹⁸ Wibawa, Febrianti, et al. "Homomorphic encryption and federated learning based privacy-preserving cnn training: Covid-19 detection use-case." *Proceedings of the 2022 European Interdisciplinary Cybersecurity Conference*. 2022.

⁸¹⁹ <https://developer.nvidia.com/blog/federated-learning-with-homomorphic-encryption/>

⁸²⁰ <https://research.ibm.com/blog/federated-learning-homomorphic-encryption>

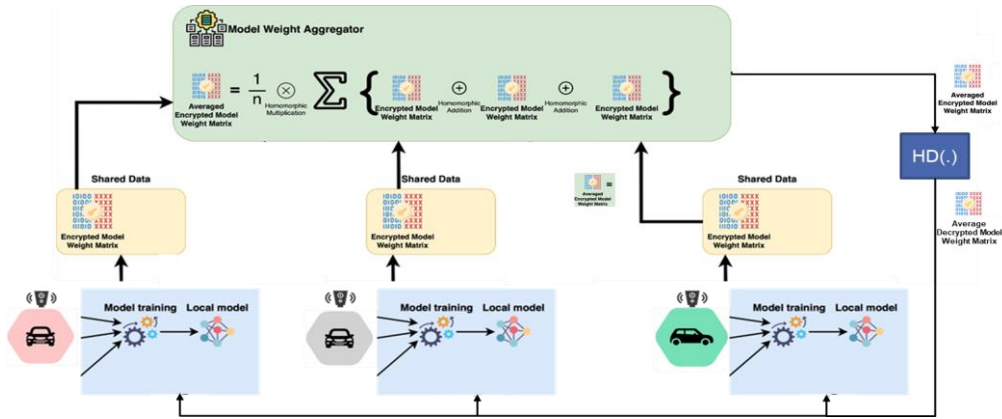


Figure 28: A generic HE-Federated learning scheme

The benefit that stems out of the above-mentioned procedure is the fact that the server does not learn the final global model or intermediate results.

FL APPLICATIONS WITHIN TRUSTEE

LEARNING BASED INTELLIGENCE FOR CONNECTED AND AUTONOMOUS VEHICLES

In the following sections, the application of FL on functionalities pertaining to connected and autonomous vehicles will be described, including (a) data generation, (b) targeted tasks (e.g., SLAM, perception) and, (c) the extension to the FL context.

EXPERIMENTAL SETUP: INTEGRATED SIMULATOR AND DATASETS

Cyber physical systems and autonomous vehicles are extremely complex systems whose behavior is practically impossible to be modelled. It is widely considered that simulation can deliver the degree of testing necessary for the development and the implementation of truly robust and secure systems.

Some of the advantages of using simulator in automotive domain are the following:

- **Safety:** Autonomous driving simulators allow developers to test their systems in a safe and controlled environment. In addition, the developer can test the model in fully controllable conditions and can generate corner cases that cannot be produced in real life scenarios. This lowers the risk of accidents and injuries, since potential safety issues can be spotted and dealt with before testing on actual roads.
- **Cost Reduction:** Real-world research and testing of autonomous driving systems can result into increased costs because they need specialized tools and a lot of resources. Yet, since simulation runs on relatively cheap hardware and does not require any physical infrastructure, it can be far more cost-effective.
- **Scalability:** It is simple to scale up autonomous driving simulations to test a variety of scenarios and circumstances, including bad weather, challenging intersections, and heavy traffic.
- **Development speed:** Simulation can be executed at accelerated speeds and requires less setup and preparation time than physical tests. As a result, system developers may test and enhance their models more quickly, resulting in more rapid progress.
- **Data generation:** Large volumes of data can be produced through simulations, and that data can be utilized to train and enhance autonomous driving systems. The analysis of this data can be utilized to find trends, improve algorithms, and create more accurate and robust autonomous driving systems.

Traffic Simulators

Traffic Simulators constitute an indispensable and quite mature tool for evaluating complex traffic systems since they can emulate the time variability of traffic phenomena⁸²¹. Two of the most popular commercial simulators are PTV Vissim⁸²², which was first launched in 1992 and Aimsun⁸²³. Both are based on microscopic simulation meaning that each entity is simulated individually, and both can provide statistical results on complex implemented traffic scenarios. SUMO⁸²⁴ is probably the most popular traffic simulator in research community since it's open source. In addition, it is compatible with formats from other simulators, like Vissim and it provides an API called TraCI which enables the bidirectional communication with other applications something that makes SUMO a suitable candidate for becoming a building block of an integrated simulator. These simulators can perform complex traffic scenarios but by themselves are not able to simulate physics events which makes them insufficient in simulating and connected autonomous vehicles' scenarios

Network Simulators

Network simulators in contrast to Traffic and Game Engine-based simulators can model and simulate properly the communications between the simulation entities. The main method utilized is simulations based on discrete events in which the simulation behaviors are not dictated by continuous equations but rather by temporal discrete events. One of the most popular simulators is ns-3⁸²⁵ which was firstly released in 2008 offers an extensive set of libraries for modeling communication channels and protocols, but still lacks in features simulation vehicular networks. OMNeT++⁸²⁶ is a free, open-source, extensible, modular, component-based C++ simulation framework which also provides an Integrated Development Environment based on Eclipse. Independent frameworks extend OMNeT++ and provide support on real-time simulation, protocol implementations and network emulation. One extension like this is Artery⁸²⁷ which it started as an extension of Veins framework but can now be used independently. Artery models and implements V2X simulations based on ETSI ITS-G5 protocols and facilitates services, such as Cooperative Awareness (CAMs) and Decentralized Notification (DENMs).

Game Engine-based Simulators

A virtual environment based on a game engine offers many comparative advantages over running experiments in the real world since there are far fewer restrictions in areas such as security, costs and control of environmental conditions. Key characteristics of a game engine like the rendering system, the physics engine and the ability to extend the functionalities of the components via scripting favor the development and execution of autonomous vehicle scenarios.

CARLA⁸²⁸ is probably the most popular game engine-based simulator amongst the research community. It is open source and based on Unreal Engine. It implements a scalable client-server architecture in which all simulation-related tasks are assigned to the server and the clients can control the simulation environment and entities via a

⁸²¹ J. Barceló, Ed., Fundamentals of Traffic Simulation. Springer, 2010. [Online]. Available: <https://EconPapers.repec.org/RePEc:spr:isorms:978-1-4419-6142-6>

⁸²² "Traffic simulation software — pvt vissim — pvt group," <https://www.ptvgroup.com/en/solutions/products/ptv-vissim/>,

⁸²³ "Aimsun: simulation and ai for intelligent mobility," <https://www.aimsun.com/>

⁸²⁴ P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic traffic simulation using sumo," in The 21st IEEE International Conference on Intelligent Transportation Systems. IEEE, 2018. [Online]. Available: <https://elib.dlr.de/124092>

⁸²⁵ "ns-3 — a discrete-event network simulator for internet systems," <https://www.nsnam.org/>,

⁸²⁶ A. Varga, OMNeT++. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 35–59. [Online]. Available: https://doi.org/10.1007/978-3-642-12331-3_3

⁸²⁷ R. Riebl. Artery architecture. [Online]. Available: <http://artery.v2x-research.eu/architecture>

⁸²⁸ A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," 2017

Python API. SVL⁸²⁹ is another end-to-end simulator platform built for the development and testing of self-driving vehicles, which is based on Unity3D⁸³⁰. Robot simulators like Gazebo⁸³¹ can also be utilized for simulating autonomous vehicles. Gazebo is a modular, open-source simulator which can use Open Dynamics Engine (ODE) as its default physics engine but can support other implementations too.

The main problem with game engine-based simulators is that they may not be of high fidelity. When a high-fidelity simulation is required, you may need to utilize an external component which contains a relevant mathematical representation to obtain realistic calculations.

Integrated Simulators

Even though, various simulation tools and frameworks exist in the area of autonomous vehicles, in many cases it is important to follow a multi-disciplinary approach, addressing the problem of simulating connected and autonomous vehicles in a holistic way covering all the aspects regarding the simulation of communications, the vehicle's behaviors, and V2X communications. Therefore, one obvious solution is the development of Integrated Simulators which combine the functionalities of tools already implemented.

A very popular combination is the coupling of a network and a traffic simulator in which both ends run the same simulation and the synchronization of their states is based on the information exchanged between the two parts. Following this approach implementations like ExNS3⁸³², Veins⁸³³ and iTETRIS⁸³⁴ implement the aforementioned paradigm. In all cases a network simulator, OMMNeT++ or ns-3, and a traffic simulator exchange information via dedicated communication modules. However, this setup lacks data generated by the physics engine and therefore is insufficient for testing scenarios that include, for example, perception algorithms.

⁸²⁹ “The most powerful real-time 3d creation tool - unreal engine,” <https://www.unrealengine.com/en-US/>, (Accessed on 02/11/2023)

⁸³⁰ “Unity real-time development platform — 3d, 2d vr & ar engine,” <https://unity.com/>, (Accessed on 02/11/2023)

⁸³¹ N. Koenig and A. Howard, “Design and use paradigms for gazebo, an open-source multi-robot simulator,” in 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), vol. 3, 2004, pp. 2149–2154 vol.3

⁸³² H. Arbabi and M. C. Weigle, “Highway mobility and vehicular ad-hoc networks in ns-3,” 2010

⁸³³ C. Sommer, R. German, and F. Dressler, “Bidirectionally Coupled Net-work and Road Traffic Simulation for Improved IVC Analysis,” IEEE Transactions on Mobile Computing (TMC), vol. 10, no. 1, pp. 3–15, January 2011.

⁸³⁴ M. Rondinone, J. Maneros, D. Krajzewicz, R. Bauza, P. Cataldi, F. Hrizi, J. Gozalvez, V. Kumar, M. Röckl, L. Lin, O. Lazaro, J. Leguay, J. Harri, S. Vaz, Y. Lopez, M. Sepulcre, M. Wetterwald, R. Blokpoel, and F. Cartolano, “itetris: A modular simulation platform for the large scale evaluation of cooperative its applications,” Simulation Modelling Practice and Theory, vol. 34, pp. 99–125, 2013. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1569190X1300018X>

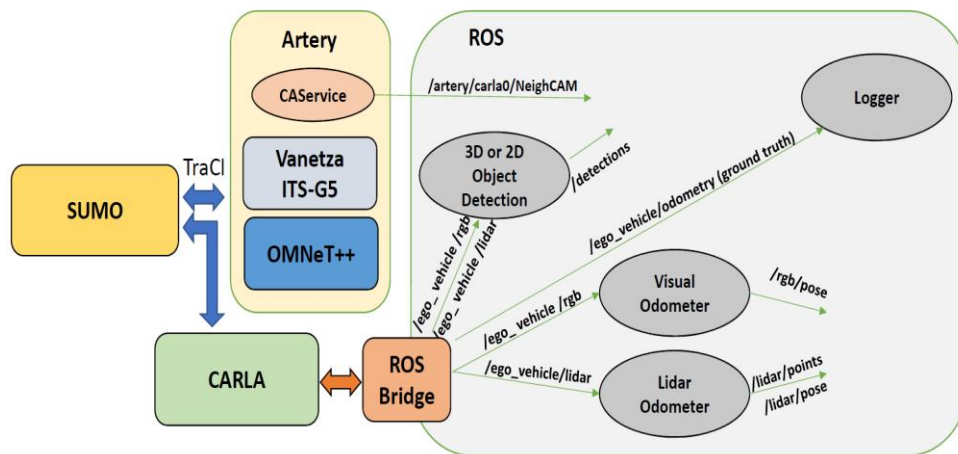


Figure 29: Integrated simulation framework

There are also works that present Integrated simulators^{835, 836} that model vehicular networks, driver models, vehicle models and control algorithms by integrating a network, a traffic, and a driving simulator. Both implementations do not have the ability to simulate physical phenomena since there is not a physics engine-based simulator. Gazebo has also been used as physics simulator⁸³⁷ in a network stack that also includes Artery for supporting cooperative scenarios such as platooning. Finally, the integrated simulation framework⁸³⁸ depicted in Figure 29, is built from available open-source components, and focuses on cooperative autonomous vehicles research and development. The framework combines the CARLA simulation engine with the ETSI ITS-G5 implementation of the Artery/OMNET++ network simulator, as well as with selected functionality available in ROS, in order to provide a complete and high-fidelity integrated simulation environment which can be exploited for the development and testing of unaltered final code running in ROS-enabled devices.

Automotive Datasets

In recent years there is a need in the automotive industry and the research community for datasets related to autonomous driving systems. Autonomous vehicles must be able to navigate complex city streets and interact safely with other vehicles, pedestrians, and objects. Moreover, they must operate properly under challenging conditions, like sudden illumination changes or poor features environments. To enable this, several large-scale datasets have been developed to aid in the development of autonomous vehicles.

⁸³⁵ G. Lee and J.-I. Jung, "Integrated simulator for evaluating cooperative eco-driving system," in 2019 International Symposium on Signal Processing and Communication Systems (ISPACS), Dec 2019, pp. 1–2.

⁸³⁶ Y. Zhao, A. Wagh, Y. Hou, K. Hulme, C. Qiao, and A. W. Sadek, "Integrated traffic-driving-networking simulator for the design of connected vehicle applications: Eco-signal case study," *Journal of Intelligent Transportation Systems*, vol. 20, no. 1, pp. 75–87, 2016, cyber Transportation Systems and Connected Vehicle Research. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1547245022000354>

⁸³⁷ B. Vieira, R. Severino, E. V. Filho, A. Koubaa, and E. Tovar, "Copadrive - a realistic simulation framework for cooperative autonomous driving applications," in 2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE), 2019, pp. 1–6

⁸³⁸ C. Anagnostopoulos, C. Koulamas, A. Lalos and C. Stylios, "Open-Source Integrated Simulation Framework for Cooperative Autonomous Vehicles," *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro, 2022, pp. 1-4, doi: 10.1109/MECO55406.2022.9797115.

KITTI⁸³⁹ can be characterized as the pioneer for this kind of datasets. It provides valuable information for challenging computer vision tasks, such as object detection and tracking, semantic and instance segmentation, visual odometry, etc. To accomplish that, there were used, a Velodyne LiDAR sensor, a GPS/IMU, and two high-resolution color and grayscale video cameras. All the sequences contain data collected from rural areas or highways during daytime. Even though KITTI was and still remains a valuable tool in research community it failed to include real world complexity.

The Cityscapes⁸⁴⁰ dataset was created five years later, to enhance the semantic understanding of urban areas. The dataset contains 5000 annotated images with fine annotations and 20,000 annotated images with coarse annotations, capturing 50 different cities during the daytime. The annotations include a variety of objects such as roads, buildings, vehicles, and pedestrians, as well as metadata from vehicle odometry. The Cityscapes dataset has been used in a variety of computer vision tasks, including semantic segmentation, object detection, and scene understanding. Another large-scale dataset for urban environments is the Mapillary Vistas dataset, which was created by G. Neuhold et al in 2017⁸⁴¹. This dataset contains 25,000 high-resolution images annotated by 66 object classes, captured under different weather and viewpoint conditions. The annotations include object instances as well as semantic segmentation labels, making it a valuable resource for tasks such as object detection and segmentation. Additionally, the nuScenes dataset, created by H. Caesar et al in 2018⁸⁴², combines lidar, camera, and radar data to provide 3D bounding boxes for 23 object classes. This dataset contains seven times more annotations and one hundred times more images than the Kitti dataset, making it a valuable resource for tasks such as 3D object detection and tracking. Fisher Yu et al generated the BDD100K⁸⁴³ dataset to address the issue of real-world complexity in driving datasets, which includes 100k images with data from various weather conditions. X. Huang et al created the ApolloScape⁸⁴⁴ dataset, which contains much richer information than previous datasets, including 100k images, 80k lidar samples, and 1000km trajectories from multiple cities under various conditions. J. Behley et al noticed the absence of a dataset designed to provide 3D scene understanding from point clouds and created the Semantic Kitti⁸⁴⁵ dataset by annotating all the sequences of the Kitti benchmark. To support startups and academic researchers, J. Geyer et al generated the A2D2⁸⁴⁶ dataset, which provides 40,000 frames with semantic segmentation image and point cloud labels and 3D bounding boxes annotations. While most previous authors did not take into account the

⁸³⁹ Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In 2012 IEEE Conference on Computer Vision and Pattern Recognition, pages 3354–3361, 2012. 1,2

⁸⁴⁰ Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3213–3223, 2016

⁸⁴¹ Gerhard Neuhold, Tobias Ollmann, Samuel Rota Buló, and ` Peter Kotschieder. The mapillary vistas dataset for semantic understanding of street scenes. In 2017 IEEE International Conference on Computer Vision (ICCV), pages 5000–5009, 2017

⁸⁴² Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nusenes: A multimodal dataset for autonomous driving. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 11618–11628, 2020

⁸⁴³ Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell. Bdd100k: A diverse driving dataset for heterogeneous multitask learning. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 2633– 2642, 2020

⁸⁴⁴ Xinyu Huang, Xinjing Cheng, Qichuan Geng, Binbin Cao, Dingfu Zhou, Peng Wang, Yuanqing Lin, and Ruigang Yang. The apolloScape dataset for autonomous driving. arXiv: 1803.06184, 2018.

⁸⁴⁵ J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall. SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences. In Proc. of the IEEE/CVF International Conf. on Computer Vision (ICCV), 2019

⁸⁴⁶ Jakob Geyer, Yohannes Kassahun, Mentar Mahmudi, Xavier Ricou, Rupesh Durgesh, Andrew S. Chung, Lorenz Hauswald, Viet Hoang Pham, Maximilian Muhlegg, Sebastian Dorn, Tiffany Fernandez, Martin Janicke, Sudesh Gana- pati Mirashi, Chiragkumar Savani, M. Sturm, Oleksandr Vorobiov, Martin Oelker, Sebastian Garreis, and Peter Schuberth. A2d2: Audi autonomous driving dataset. ArXiv, abs/2004.06320, 2020

influence of HD maps for tracking and motion prediction in AV-related applications, M. Chang et al provided the Argoverse⁸⁴⁷ dataset, including HD maps with semantic metadata to improve the robustness of perception mechanisms. J. Houston and R. Kesten et al contributed the Lyft⁸⁴⁸ dataset, which consists of over 1000 hours of movement of traffic agents along with their 3D bounding boxes, designed for perception and prediction. P. Sun and S. Ettinger et al developed the Waymo dataset⁸⁴⁹, which includes multiple 2D/3D labels, object trajectories, and 3D maps. Y. Liao et al created the Kitti-360 dataset⁸⁵⁰, a successor to the Kitti dataset that contains 300k images and point clouds with 2D/3D semantics of a suburban area. WoodScape⁸⁵¹, created by S. Yogamani et al, provides 2D/3D bounding boxes, semantic segmentation, soiling detection, odometry data, and other metadata to adapt deep learning algorithms for fisheye cameras. Lastly, J. Mao et al provided the ONCE⁸⁵² dataset, which is twenty times longer than the nuScenes or Waymo, consisting of 1 million point clouds and 7 million images, capturing data for 144 hours. The work by Wenzel et al. focuses on demanding perceptual conditions in automotive scenarios. Finally, CarlaScenes⁸⁵³ is a large-scale simulation dataset captured using the CARLA simulator. The dataset is oriented to address the challenging odometry scenarios that cause the current state of art odometers to deviate from their normal operations. Based on a case study of failures presented in experiments, the dataset focuses on 7 different sequences of data that implement different scenarios.

DEEP LEARNING BASED MULTI-MODAL AUTOMOTIVE SCENE ANALYSIS

⁸⁴⁷ Ming-Fang Chang, John Lambert, Patsorn Sangkloy, Jagjeet Singh, Sławomir Bak, Andrew Hartnett, De Wang, Peter Carr, Simon Lucey, Deva Ramanan, and James Hays. Argoverse: 3d tracking and forecasting with rich maps. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 8740–8749, 2019.

⁸⁴⁸ J. Houston, G. Zuidhof, L. Bergamini, Y. Ye, A. Jain, S. Omari, V. Iglovikov, and P. Ondruska. One thousand and one hours: Self-driving motion prediction dataset. <https://level-5.global/level5/data/>, 2020

⁸⁴⁹ Scott M. Ettinger, Shuyang Cheng, Benjamin Caine, Chenxi Liu, Han Zhao, Sabeek Pradhan, Yuning Chai, Benjamin Sapp, C. Qi, Yin Zhou, Zoey Yang, Aurelien Chouard, Pei Sun, Jiquan Ngiam, Vijay Vasudevan, Alexander McCauley, Jonathon Shlens, and Drago Anguelov. Large scale interactive motion forecasting for autonomous driving : The waymo open motion dataset. ArXiv, abs/2104.10133, 2021

⁸⁵⁰ Yiyi Liao, Jun Xie, and Andreas Geiger. Kitti-360: A novel dataset and benchmarks for urban scene understanding in 2d and 3d. ArXiv, abs/2109.13410, 2021

⁸⁵¹ Senthil Kumar Yogamani, Ciaran Hughes, Jonathan Hor- ´ gan, Ganesh Sistu, Pdraig Varley, Derek O’Dea, Michal U´ria´r, Stefan Milz, Martin Simon, Karl Amende, Christian Witt, Hazem Rashed, Sumanth Chennupati, Sanjaya Nayak, Saquib Mansoor, Xavier Perroton, and Patrick Perez. Wood- ´ scape: A multi-task, multi-camera fisheye dataset for autonomous driving. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pages 9307–9317, 2019

⁸⁵² Jiageng Mao, Minzhe Niu, Chenhan Jiang, Hanxue Liang, Xiaodan Liang, Yamin Li, Chao Ye, Wei Zhang, Zhenguo Li, Jie Yu, Hang Xu, and Chunjing Xu. One million scenes for autonomous driving: Once dataset. ArXiv, abs/2106.11037, 2021

⁸⁵³ A. Kloukiniotis *et al.*, "CarlaScenes: A synthetic dataset for odometry in autonomous driving," 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA, 2022, pp. 4519-4527, doi: 10.1109/CVPRW56347.2022.00498.

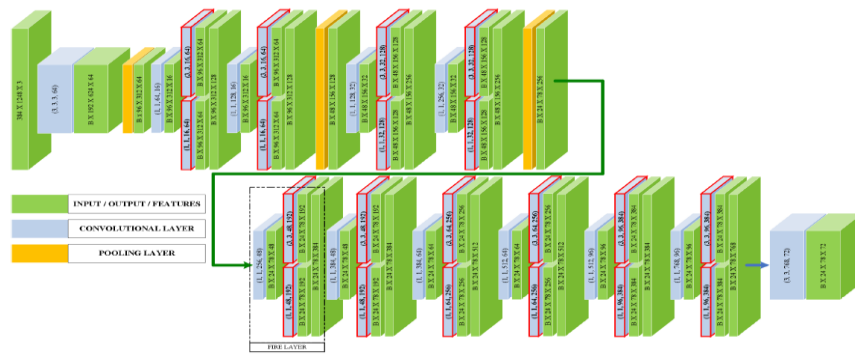


Figure 30: The architecture of SqueezeDet

Object detection has evolved considerably since the appearance of deep convolutional neural networks⁸⁵⁴. Nowadays, there are two main branches of proposed techniques in image-based object detection, namely two-stage and single stage detectors. In the first one, the object detectors, using two stages, generate region proposals which are subsequently classified in the categories that are determined by the application at hand (e.g., vehicles, cyclists and pedestrians, in the case of autonomous driving). Some important, representative, high-performance examples of this first branch are R-CNN⁸⁵⁵, Fast R-CNN⁸⁵⁶, Faster R-CNN⁸⁵⁷, Spatial Pyramid Pooling Net⁸⁵⁸, Region-based Fully Convolutional Network (R-FCN)⁸⁵⁹, Feature Pyramid Network (FPN)⁸⁶⁰ and Mask R-CNN⁸⁶¹. In the second branch, object detection is cast to a single-stage, regression-like task with the aim to provide directly both the locations and the categories of the detected objects. Notable examples, here, are Single Shot MultiBox Detector

⁸⁵⁴ Z. -Q. Zhao, P. Zheng, S. -T. Xu and X. Wu, "Object Detection With Deep Learning: A Review," in IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11, pp. 3212-3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

⁸⁵⁵ R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 580-587, doi: 10.1109/CVPR.2014.81.

⁸⁵⁶ R. Girshick, "Fast R-CNN," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

⁸⁵⁷ S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.

⁸⁵⁸ K. He, X. Zhang, S. Ren and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1904-1916, 1 Sept. 2015, doi: 10.1109/TPAMI.2015.2389824.

⁸⁵⁹ J. Dai, Y. Li, K. He, J. Sun, R-fcn: Object detection via region-based fully convolutional networks, arXiv preprint arXiv:1605.06409 (2016).

⁸⁶⁰ T. -Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 936-944, doi: 10.1109/CVPR.2017.106.

⁸⁶¹ K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2980-2988, doi: 10.1109/ICCV.2017.322.

(SSD)⁸⁶², SqueezeDet⁸⁶³, YOLO⁸⁶⁴, YOLOv3⁸⁶⁵ and EfficientDet⁸⁶⁶. Finally, accelerating deep neural networks for efficient scene analysis and understanding in automotive applications is discussed in^{867,868}.

Object detection in LIDAR point clouds is a three-dimensional problem where the sampled points are not uniformly distributed over the objects in the scene and do not directly correspond to a cartesian grid. 3D object detection is dominantly performed with 3D convolutional networks due to the irregularity and lack of apparent structure in the point cloud. Several transformations take place to match the point cloud to feature maps that are forwarded into deep networks. Commendable detection outcomes appear in the literature as early as 2016. For example, the authors of ⁸⁶⁹ projected the 3D points in a 2D map and employed 2D fully convolutional networks to successfully detect cars in a LIDAR point cloud reaching an accuracy of **71.0%** for car detection of moderate difficulty. A follow-up paper⁸⁷⁰ proposes 3D fully convolutional networks reporting accuracy of **75.3%** for car detection of moderate difficulty. However, since dense 3D fully convolutional networks demonstrate high execution times, the authors of⁸⁷¹ investigated an improved sparse convolution method for such networks, which significantly increases the

⁸⁶² Liu, W. et al. (2016). SSD: Single Shot MultiBox Detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science(), vol 9905. Springer, Cham. https://doi.org/10.1007/978-3-319-46448-0_2.

⁸⁶³ B. Wu, A. Wan, F. Iandola, P. H. Jin and K. Keutzer, "SqueezeDet: Unified, Small, Low Power Fully Convolutional Neural Networks for Real-Time Object Detection for Autonomous Driving," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017, pp. 446-454, doi: 10.1109/CVPRW.2017.60.

⁸⁶⁴ J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

⁸⁶⁵ J. Redmon, A. Farhadi, Yolov3: An incremental improvement, arXiv preprint arXiv:1804.02767 (2018).

⁸⁶⁶ M. Tan, R. Pang and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 10778-10787, doi: 10.1109/CVPR42600.2020.01079.

⁸⁶⁷ E. V. Pikoulis, C. Mavrokefalidis, S. Nousias, A. S. Lalos, A New Clustering-Based Technique for the Acceleration of Deep Convolutional Networks, Springer Singapore, Singapore, 2022, pp. 123–150. URL: https://doi.org/10.1007/978-981-16-3357-7_5. doi:10.1007/978-981-16-3357-7_5.

⁸⁶⁸ S. Nousias, E. V. Pikoulis, C. Mavrokefalidis, A. S. Lalos, Accelerating deep neural networks for efficient scene understanding in automotive cyber-physical systems, in: 2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS), 2021, pp. 63–69. doi:10.1109/ICPS49255.2021.9468126.

⁸⁶⁹ B. Li, T. Zhang, T. Xia, Vehicle detection from 3d lidar using fully convolutional network, arXiv preprint arXiv:1608.07916 (2016).

⁸⁷⁰ B. Li, 3d fully convolutional network for vehicle detection in point cloud, in: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2017, pp. 1513–1518.

⁸⁷¹ Y. Yan, Y. Mao, B. Li, Second: Sparsely embedded convolutional detection, Sensors 18 (2018) 3337.

speed of both training and inference. According to KITTI benchmarks, the reported accuracy reaches **78.6%** for

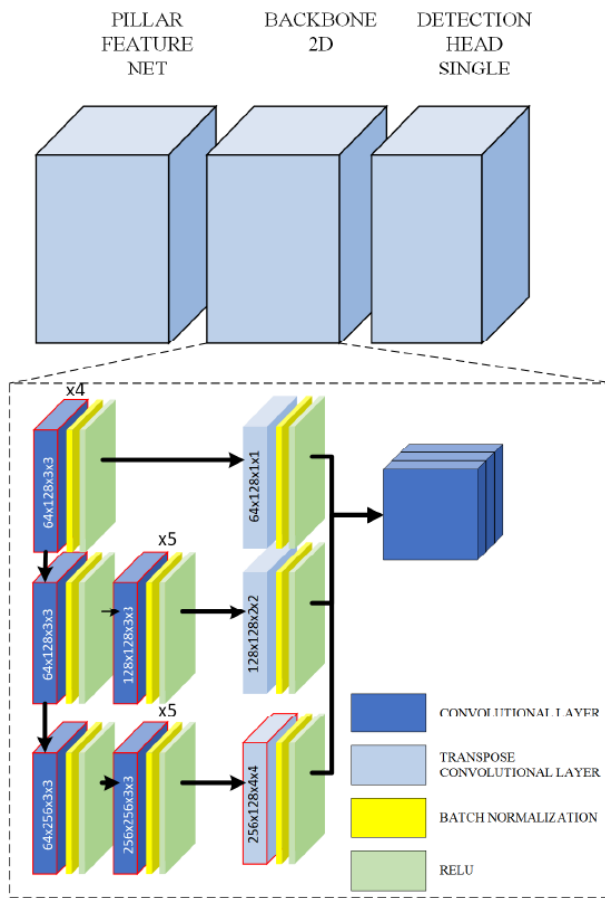


Figure 31: Architecture of PointPillars

car detection of moderate difficulty. To revisit 2D convolutions in 3D object detection, the authors in ⁸⁷² proposed Pointpillars, a novel encoder that utilizes PointNets to learn a representation of point clouds organized in vertical columns (pillars) and subsequently employed a series of 2D convolutions. Pointpillars reported accuracy of **77.28%** in the same category. The authors of ⁸⁷³ proposed PointRCNN for 3D object detection from raw point cloud. They devised a two stages stage approach where the first stage yields a bottom-up 3D proposal generation the second stage refines the proposals in the canonical coordinates to obtain the final detection results, reporting accuracy of **78.70%**. An extended variation of PointRCNN is the part-aware and aggregation neural network. The part-aware stage for the first time fully utilizes free-of-charge part supervisions derived from 3D ground-truth boxes to simultaneously predict high-quality 3D proposals and accurate intra-object part locations. Then the part-aggregation stage learns to re-score the box and refines the box location by exploring the spatial relationship of the pooled intra-object part locations. The reported accuracy reaches **79.40%**. Finally, the authors of ⁸⁷⁴ designed a compressed version of Pointpillars for automotive scene analysis, with a negligible loss in 3D detection accuracy.

⁸⁷² A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, O. Beijbom, Pointpillars: Fast encoders for object detection from point clouds, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 12697–12705.

⁸⁷³ S. Shi, X. Wang, H. Li, Pointrcnn: 3d object proposal generation and detection from point cloud, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 770–779.

⁸⁷⁴ S. Nousias, E.-V. Pikoulis, C. Mavrokefalidis, A. S. Lalos, K. Moustakas, Accelerating 3d scene analysis for autonomous driving on embedded ai computing platforms, in: 2021 IFIP/IEEE 29th International Conference on Very Large Scale Integration (VLSI-SoC), 2021, pp. 1–6. doi:10.1109/VLSI-SoC53125.2021.9606990.

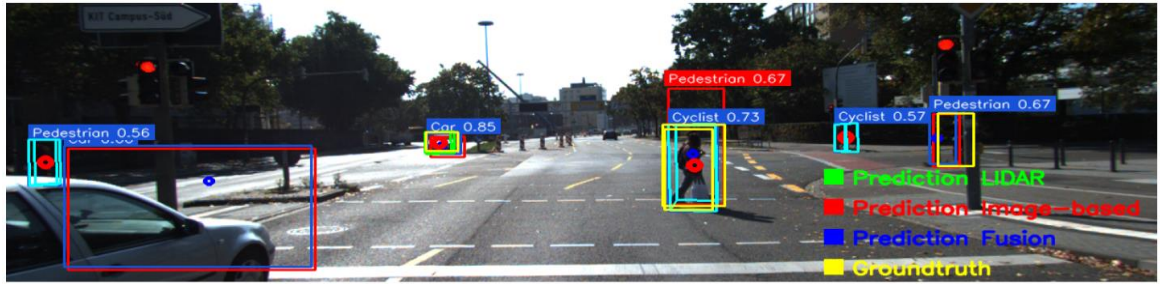
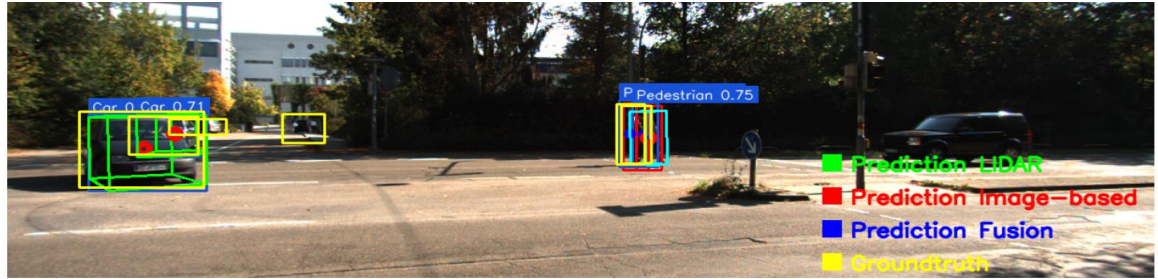


Figure 32: An example of object detection when the object detectors are LIDAR-based, Image-based and Fusion-based



Object detection from a single the point of view of a single agent is definitely vulnerable to a series of sensor limitations that can significantly affect the outcome. These limitations entail occlusion, limited field-of-view and low-point density at distant regions. The transition from isolated CPSs to CPSoSs enabling the collaboration among various agents offers the opportunity to tackle such problems. The authors of ⁸⁷⁵ proposed a cooperative sensing scheme where each CPS combines its sensing data with that of other connected vehicles to help enhance perception. Furthermore, to tackle the increased amount of data, the authors propose a Sparse Point-cloud Object Detection. It is important to highlight that the agents share on-board V2V information and fuse locally. The authors of ⁸⁷⁶ examined feature-level fusion in a follow-up work. The authors proposed F-Cooper framework, a method that improves the autonomous vehicle's detection precision without introducing much computational overhead aiming to utilize the capacity of feature maps especially for 3D LiDAR data generated on autonomous vehicles, as the feature maps are used for object detection only by single vehicles. F-Cooper is an end-to-end 3D object detection with feature-level fusion supporting voxel feature fusion and spatial feature fusion. Voxel feature fusion achieves almost the same detection precision improvement when compared to the raw-data level fusion solution, the latter offers the ability to dynamically adjust the size of feature maps to be transmitted. A unique characteristic of F-Cooper is that it can be deployed and executed on in-vehicle and roadside edge systems. The authors of ⁸⁷⁷ proposed TruPercept to tackle malicious attacks against cooperative perception systems. In their trust scheme, each agent reevaluates the detection originating from its neighbouring agents using data from its position and perspective. The authors of ⁸⁷⁸ proposed a centralized system that fuses data from multiple infrastructure sensors which facilitate

⁸⁷⁵ Q. Chen, S. Tang, Q. Yang, S. Fu, Cooper: Cooperative perception for connected autonomous vehicles based on 3d point clouds, in: 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), IEEE, 2019a, pp. 514–524.

⁸⁷⁶ Q. Chen, X. Ma, S. Tang, J. Guo, Q. Yang, S. Fu, F-cooper: Feature based cooperative perception for autonomous vehicle edge computing system using 3d point clouds, in: Proceedings of the 4th ACM/IEEE Symposium on Edge Computing, 2019b, pp. 88–100.

⁸⁷⁷ B. Hurl, R. Cohen, K. Czarnecki, S. Waslander, Trupercept: Trust modelling for autonomous vehicle cooperative perception from synthetic data, in: 2020 IEEE Intelligent Vehicles Symposium (IV), IEEE, 2020, pp. 341–347.

⁸⁷⁸ E. Arnold, M. Dianati, R. de Temple, S. Fallah, Cooperative perception for 3d object detection in driving scenarios using infrastructure sensors, IEEE Transactions on Intelligent Transportation Systems (2020).

the management of both sensor and processing costs through shared resources. The authors deployed VoxelNet⁸⁷⁹ and claim to have reached average precision score $AP_{3D} = 98\%$ for early fusion strategy in and $AP_{3D} = 81\%$ for late fusion strategy in a T-Junction. In a more recent study⁸⁸⁰, a Cooperative Spatial Feature Fusion (CoFF) to address F-Cooper limitations was proposed. F-Cooper's *maxout* function treats similarly feature maps from different sources and conditions leading to mis-classifications. CoFF calculates and assigns importance weights to the received feature maps given the data available to the ego-vehicle. The authors claim to reach **90%** improvement in average precision for far object cases with respect to F-Cooper.

The high performance of deep neural networks (DNNs) is typically related to analogously high requirements regarding computational and storage resources. This becomes problematic in automotive applications due to the necessity of very fast inference times on the one hand, and the limited computational, storage, and energy resources of mobile systems, on the other. To allow for high-performing DNN models to be deployed on systems with limited resources (e.g. mobile embedded devices), MCA has become recently an active area of research⁸⁸¹. MCA comprises techniques for transforming existing, pre-trained, highly performing yet resource demanding DNN models, into lighter versions while mitigating the impact on the achieved performance. Although there are many algorithms for MCA, three main directions can be identified. Many techniques remove parts of the networks by proposing metrics for assessing their importance (e.g., removing filters⁸⁸²). In another direction, the representation of the involved parameters is limited by reducing the number of bits for each scalar value, or by grouping parameters and substituting them with appropriately selected representatives using vector and product quantization⁸⁸³, leading to the so-called weight sharing techniques. Finally, some techniques exploit mathematical properties of the tensors describing the layers of the networks like sparsity and low rankness to achieve the desired MCA gains⁸⁸⁴.

DEEP LEARNING BASED MULTI-MODAL ODOMETERS AND SLAM

Simultaneous localization and mapping (SLAM) is very important for the localization and the achievement or persistent autonomy for autonomous vehicles especially in environments where there is no prior knowledge of the environment. SLAM systems have gained extreme popularity over the last years and have made great progress on providing accurate systems that can localize autonomous agent. However, they still face considerable issues when they face great illumination changes and are deployed in large-scale environments. Therefore, the requirement for more robust systems and semantic perception drives the research to non-model-based techniques. Deep learning-based techniques seem like the logical evolution of SLAM systems since they can provide more robust and high-level features. In the following sections, after a short reference to traditional SLAM systems, we present the three of the most popular types of NN used in SLAM systems, and then we describe the current state of the art in deep based algorithms concerning image feature extraction and matching, depth estimation, lidar odometry, and sensor fusion.

⁸⁷⁹ Y. Zhou, O. Tuzel, Voxelnet: End-to-end learning for point cloud based 3d object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4490–4499.

⁸⁸⁰ J. Guo, D. Carrillo, S. Tang, Q. Chen, Q. Yang, S. Fu, X. Wang, N. Wang, P. Palacharla, Coff: Cooperative spatial feature fusion for 3d object detection on autonomous vehicles, IEEE Internet of Things Journal (2021).

⁸⁸¹ L. Deng, G. Li, S. Han, L. Shi, and Y. Xie, “Model compression and hardware acceleration for neural networks: A comprehensive survey,” Proceedings of the IEEE, vol. 108, no. 4, pp. 485–532, 2020.

⁸⁸² S. Lin, R. Ji, Y. Li, C. Deng, and X. Li, “Toward compact convnets via structure-sparsity regularized filter pruning,” IEEE transactions on neural networks and learning systems, 2019.

⁸⁸³ Y. Gong, L. Liu, M. Yang, and L. Bourdev, “Compressing deep convolutional networks using vector quantization,” arXiv:1412.6115, 2014.

⁸⁸⁴ S. Bhattacharya and N. D. Lane, “Sparsification and separation of deep learning layers for constrained resource inference on wearables,” in Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM, 2016, pp. 176–189.

Traditional SLAM

SLAM is a very active area of research in scientific domains such as robotics and automotive. The most notable sensors to be used for SLAM are camera and LIDAR. The principle behind the two approaches is actually the same: determine the motion of sensor as well as construct the surrounding map of vehicle given the movement of pixels or 3D points in frame sequences. Typically, a front-end module roughly estimates the pose and the depth of landmarks. Subsequently, the back-end module performs a global windowed optimization for the determination of final pose and reconstruction of the map. Considering the case of camera sensor, two main approaches can be identified: i) feature based, which rely on salient points (or features) extraction, feature matching (or registration) between successive image frames and finally camera motion estimation through n-points algorithm, ($n \in \{4, 6, 8\}$ depending on motion's degrees of freedom), ii) direct based, which use the entire image instead of extracting time-consuming interesting points. The registration is based on photometric error consistency, instead of the geometric position of pixels as in feature-based solutions. SLAM realized by RGB-D cameras reconstructs highly accurate the surrounding map due to direct depth estimation, however it performs well only for indoor scenarios since it is very sensitive to daylight. Popular methods in this camera-based category are Direct Sparse Odometry (DSO) for rectangular images⁸⁸⁵, fisheye data⁸⁸⁶ and loop closure task⁸⁸⁷, as well as ORB-SLAM⁸⁸⁸. Camera based SLAM is usually prone to errors related to sensitivity to texture and light, photometric calibration, etc. LIDAR on the contrary, when applied to the SLAM problem, can achieve a low drift motion estimation for an acceptable computational complexity. The main solution for LIDAR based navigation is the scan-matching approach followed by a graph optimization. Scan-matching is the fundamental process used to create 3D maps with LIDAR and gives precise information on motion. The general approach for registering 3D point clouds is based on the ICP⁸⁸⁹. Lightweight and Ground-Optimized Lidar Odometry and Mapping (LeGO-LOAM) is a well known LIDAR-based SLAM method⁸⁹⁰, which adopts a feature based concept. More specifically, the point cloud is projected onto a range image. Segmentation is applied then, in order to remove clusters of points with size less than 30. Afterwards, feature extraction is followed, identifying edge and planar features in the image domain. Finally, Levenberg-Marquardt optimization is applied to the matched feature points for estimating the pose. The main challenge of LIDAR-based solutions is the very high storage requirements but also the tradeoff between economic cost and accuracy. The cost of the much more accurate 64-channel LIDAR is prohibitive for realistic commercial deployment with respect to less accurate 16-channel LIDAR. Moreover, the registration of point-clouds is often failed for cluttered outdoor scenes as well as serious degradation in performance is noticed when we experience scattering effect under rain, snow, or fog, etc.

Popular Types of Neural Networks used in SLAM systems.

Convolutional Neural Networks (CNN) can be characterized as one of the most popular deep neural network architectures, especially in the case of computer vision applications. It consists mainly of vision layers (convolutional, activation and pooling layers) and fully connected layers (Figure 33). However, drop out and

⁸⁸⁵ J. Engel, V. Koltun and D. Cremers, "Direct Sparse Odometry," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 3, pp. 611-625, 1 March 2018, doi: 10.1109/TPAMI.2017.2658577.

⁸⁸⁶ H. Matsuki, L. von Stumberg, V. Usenko, J. Stückler, and D. Cremers, "Omnidirectional dso: Direct sparse odometry with fisheye cameras," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3693-3700, 2018

⁸⁸⁷ X. Gao, R. Wang, N. Demmel, and D. Cremers, "LDSO: Direct Sparse Odometry with Loop Closure." *arXiv*, 2018. doi: 10.48550/ARXIV.1808.01111.

⁸⁸⁸ R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "Orb-slam: A versatile and accurate monocular slam system," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147-1163, 2015

⁸⁸⁹ P. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239-256, 1992

⁸⁹⁰ T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4758-4765

normalization layers are also frequently used. The choice of the loss function, which determines the distance between the prediction and the ground truth is very vital for CNN and affects the convergence of the model. Cross-entropy is usually used in classification problems and Euclidean in regression. Even though CNNs can capture the spatial information in a picture, they are unable to perceive the temporal dimension.

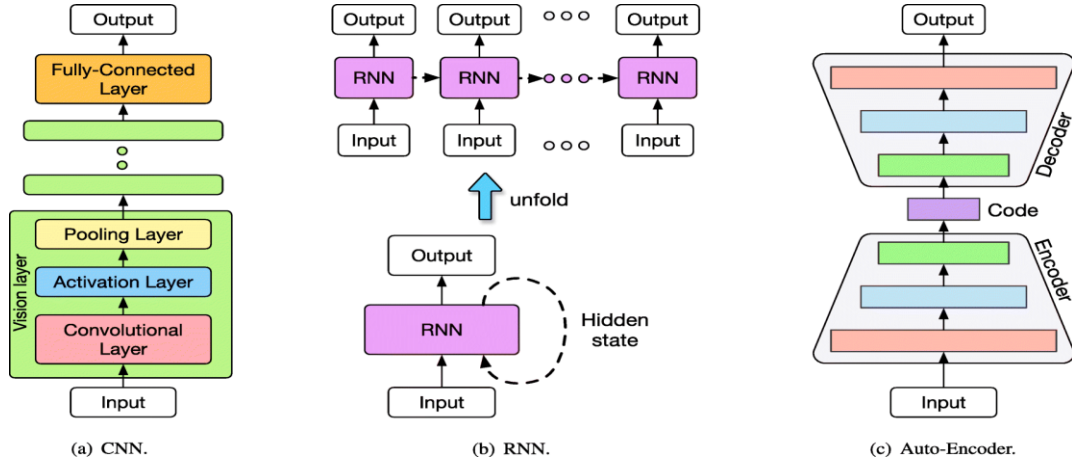


Figure 33: Popular types NNs used in SLAM systems⁸⁹¹

The Recurrent Neural Networks (RNN) are designed to deal with data that encapsulate temporal information, like in the case of video clips. They utilize feedback loops to maintain the memory of their hidden state over time (Figure 33b). if x_t , y_t and h_t are an input state, an output state and a hidden state at time instance t , the new updated hidden and the output state of a simple RNN can be calculated by the following equation:

$$h_t = H(W_{xh}x_t + W_{hh}h_{t-1} + b_h)$$

$$y_t = W_{hy}h_t + b_y$$

Where h_t and y_t are the hidden and output state at time instance t , W is the Weight matrix, b is the bias, and $H()$ is a non-linear function.

The long short-term memory (LSTM) architecture comes to address one of the biggest issues faced by the RNN, the vanishing gradient problem. LSTMs use several gates in order to control what is forgotten or not by the network. They use a different repeating model, in which the information propagated to the next module, is controlled by these gates.

Autoencoders (Figure 33c) can be categorized as a special case of CNN. They are built to perform three distinct operations: encoding, coding and decoding. Their main purpose is the transformation of an input to a different representation. More specifically, the autoencoder firstly maps the input x to an intermediate representation y

$$y = e(x)$$

where $e()$ is the non-linear function representing the operation of the encoder. Then the decoder transforms the intermediate state y to the output z , which is usually a representation that contains the same main features with the input x . This operation is illustrated in the following equation:

⁸⁹¹ R. Li, S. Wang and D. Gu, "Ongoing Evolution of Visual SLAM from Geometry to Deep Learning: Challenges and Opportunities," *Cognitive Computation*, vol. 10, p. 875–889, September 2018.

$$\mathbf{z} = d(\mathbf{y})$$

where $d()$ is the non-linear operation of the decoder. The main areas of application of autoencoders in Visual Odometry Systems are depth estimation, feature extraction and semantic segmentation.

Feature Extraction

For a VO to be reliable and accurate, robust feature points must be extracted. The essence behind the feature point extraction is the selection of a reliable and robust set of features that is repeatable and detectable in the subsequent frames. The tracking process includes data association, thus the feature matching over time. The vSLAM approaches that are presented below suggest various enhancements to feature point selection and data association.

Quad Networks⁸⁹² utilized unsupervised learning for the training of a network for producing feature points. The network is trained to detect features that are repeatable in warped images. However, this implementation can be applied only on patches and doesn't create descriptors.

SuperPoint⁸⁹³ uses a common VGG based encoder for producing a common representation which is then fed to two parallel branches that produce simultaneously pixel-level locations of interest points and the 256-bit long associated descriptors in one forward pass. The proposed framework is self-supervised, and it doesn't extract patches, but it rather operates on full-sized images. It also introduces the stage of Homographic Adaption in which pseudo ground truth labels are produced by a network trained on online-generated synthetic images of basic geometric shapes. Then, the annotated dataset is used for training the final model which is based on Siamese training scheme.

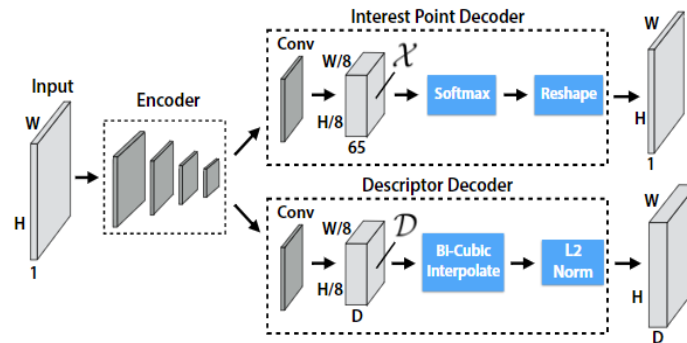


Figure 34: Superpoint architecture

UnsuperPoint⁸⁹⁴ adopts parts of the Superpoint architecture such as the shared computation of the detector and the descriptor and presents a self-supervised deep learning-based interest point detector. It introduces a new loss function that uses regression of point positions. It uses a shared backbone segment, which down samples an RGB image.

⁸⁹² Nikolay Savinov, Akihito Seki, Lubor Ladicky, Torsten Sattler, and Marc Pollefeys. Quad-networks: unsupervised learning to rank for interest point detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1822–1830, 2017.

⁸⁹³ D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperPoint: Self-Supervised Interest Point Detection and Description,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2018. doi: 10.1109/cvprw.2018.00060.

⁸⁹⁴ P. H. Christiansen, M. F. Kragh, Y. Brodskiy, and H. Karstoft, “UnsuperPoint: End-to-end Unsupervised Interest Point Detector and Descriptor.” arXiv, 2019. doi: 10.48550/ARXIV.1907.04011..

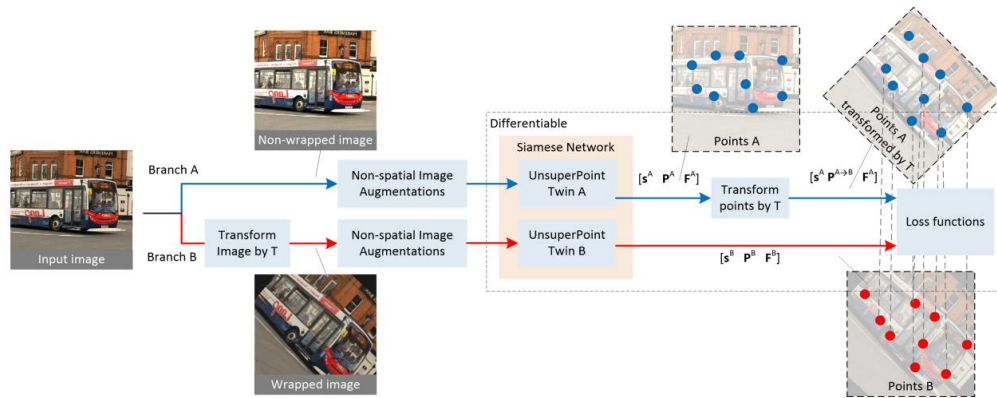


Figure 35: Siamese training used in UnsuperPoint. The original image and the warped are forwarded through the network. Correspondences between the two images are used for training the model (Christiansen, Kragh, Brodskiy, & Karstoft, 2019)

Another work that simultaneously computes the interest points and the relevant descriptors, thus following the describe-and-detect approach, is D2-NET⁸⁹⁵. In contrast to Superpoint, where there are two different decoders for detections and descriptors, in D2-NET, detection and description share a common set of parameters and undergo the same joint optimization process. The VGG-16 based convolutional neural network for feature extraction is trained on real-world annotated data.

The Repeatable and Reliable Detector and Descriptor, R2D2⁸⁹⁶, proposes a novel unsupervised network which produces three outputs for an image. The first output is a 3D tensor which corresponds to D-size descriptors for each pixel. The second output is a heatmap which gives sparse and repeatable interest point locations. The last output is an associated reliability map that illustrates the estimated reliability. The network was trained in both synthetic and real data using style transfer techniques to address issues like drastic illumination changes. In the work of Tyszkiewicz¹ et al⁸⁹⁷ is described a method for learning local features with policy gradient. More, specifically, a backbone network takes images as input and produces keypoint heatmaps and dense descriptors. Simple points are selected by these heatmaps by sampling and the descriptors that correspond to these points are used to create a distribution over feature matches across images. Using Reinforcement Learning principles and the ground truth, they assign positive or negative rewards to each match and finally the execute gradient descent to maximize the anticipated reward. ContextDesc⁸⁹⁸ aggregates the cross-modality contextual information by including geometric context from the distribution of the keypoints in the 2d space and visual context from high level feature representations.

LIFT⁸⁹⁹ proposes an end-to-end differentiable framework which consists of three major components that feed to each other: the Detector, the Orientation Estimator, and the Descriptor. Each component is based on a CNN and they are meshed together with the use of Spatial Transformers. LIFT uses a Siamese network for training but it

⁸⁹⁵ M. Dusmanu et al., “D2-Net: A Trainable CNN for Joint Detection and Description of Local Features.” arXiv, 2019. doi: 10.48550/ARXIV.1905.03561

⁸⁹⁶ J. Revaud et al., “R2D2: Repeatable and Reliable Detector and Descriptor.” arXiv, 2019. doi: 10.48550/ARXIV.1906.06195.

⁸⁹⁷ M. J. Tyszkiewicz, P. Fua, and E. Trulls, “DISK: Learning local features with policy gradient.” arXiv, 2020. doi: 10.48550/ARXIV.2006.13566

⁸⁹⁸ Z. Luo et al., “ContextDesc: Local Descriptor Augmentation with Cross-Modality Context.” arXiv, 2019. doi: 10.48550/ARXIV.1904.04084.

⁸⁹⁹ H. M. S. Bruno and E. L. Colombari, *LIFT-SLAM: A deep-learning feature-based monocular visual SLAM method*, vol. 455. Elsevier BV, 2021, pp. 97–110. doi: 10.1016/j.neucom.2021.05.027

receives as input feature points produced by a Structure-from-Motion (SfM) algorithm. LF-NeT⁹⁰⁰, like LIFT, selects suitable patches of the image it transforms them using a Spatial Transform Network. Then it feeds the representation into the descriptor module. The framework is differentiable end-to-end, and it can predict translation, rotation, scale and descriptors.

Matching

Deep learning-based matching methods are mainly used to substitute traditional algorithms in information extraction and representation or model regression⁹⁰¹. Point matching aims at finding the features or pixel correspondences between two images using or not using feature detection techniques and descriptors. All the relevant techniques can roughly fall under two big categories: area matching techniques and feature matching techniques.

Area matching techniques usually employ CNNs for image-level information extraction and similarity measurement. The best way to achieve area-based image registration is to acquire an initial estimation of a similarity metric between the pair of images and use this as a trigger for an iterative optimization procedure. Thus, deep learning-based metrics are used in place of traditional metrics. For example, Revaud et al.⁹⁰² proposed Deepmatching, which is an algorithm that generates dense correspondences between two images, utilizing a multi-scale approach on scoring the matches and handling non-rigid deformations. Arar et al.⁹⁰³ introduced an unsupervised method for multi-modal registration using two networks, a spatial transformation network and a translation network. Sokooti et al.⁹⁰⁴ avoid using a dissimilarity metric and rather employed an end-to-end CNN architecture that directly estimates the displacement vector field of pair of images. Similarly, Chen et al.⁹⁰⁵ presented ARvicon-Net, an end-to-end trainable deep neural network which directly predicts the dense displacement field for pairwise image alignment. The Global-Local Universal Network for Dense Flow and Correspondences (GLU-Net)⁹⁰⁶ is a solution that tackles dense correspondence problems in a uniform way.

Regarding feature matching techniques they can be further divided into two categories, techniques which use feature descriptors and techniques that don't. A prominent example of the first category is SuperGlue⁹⁰⁷, which is a detector implemented by the same research team that presented Superpoint, receives as input a pair of interest points and descriptors that correspond to two images. The architecture of SuperGlue (Figure 36) can be roughly split into two

⁹⁰⁰ Nikolay Savinov, Akihito Seki, Lubor Ladicky, Torsten Sattler, and Marc Pollefeys. Quad-networks: unsupervised learning to rank for interest point detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1822–1830, 2017

⁹⁰¹ Ma, J., Jiang, X., Fan, A. et al. Image Matching from Handcrafted to Deep Features: A Survey. *Int J Comput Vis* **129**, 23–79 (2021). <https://doi.org/10.1007/s11263-020-01359-2>

⁹⁰² Revaud, J., Weinzaepfel, P., Harchaoui, Z., & Schmid, C. (2016). Deepmatching: Hierarchical deformable dense matching. *International Journal of Computer Vision*, *120*(3), 300–32

⁹⁰³ Arar, M., Ginger, Y., Danon, D., Bermano, A. H., & Cohen-Or, D. (2020). Unsupervised multi-modal image registration via geometry preserving image-to-image translation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 13,410–13,419.

⁹⁰⁴ Sokooti, H., de Vos, B., Berendsen, F., Lelieveldt, B. P., Isgum, I., & Staring, M. (2017). Nonrigid image registration using multi-scale 3d convolutional neural networks. In *Proceedings of the international conference on medical image computing and computer-assisted intervention*, pp. 232–239

⁹⁰⁵ Chen, J., Wang, L., Li, X., & Fang, Y. (2019). Arbicon-net: Arbitrary continuous geometric transformation networks for image registration. In *Advances in neural information processing systems*, pp. 3410–3420.

⁹⁰⁶ Truong, P., Danelljan, M., & Timofte, R. (2020). Glu-net: Global-local universal network for dense flow and correspondences. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 6258–6268.

⁹⁰⁷ P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperGlue: Learning Feature Matching With Graph Neural Networks,” in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2020. doi: 10.1109/cvpr42600.2020.00499.

components. The Attentional Graph Neural Network receives as input the encoded key point positions and their descriptors in a single vector for each image. It is composed of a set of L alternating self-attention and cross-attention layers. The outcome of this components are strong feature representations, which are fed to the Optimal Matching Layer. This component creates a score matrix, augments it with dustbins, and finds the optimal matches between the points.

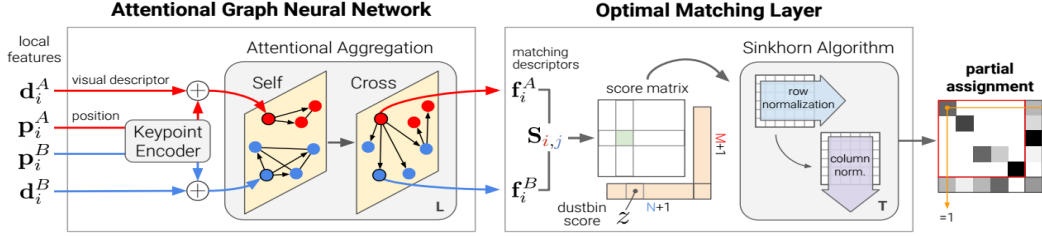


Figure 36: Overview of SuperGlue

On the other hand, LoFTR⁹⁰⁸ uses transformers for accomplishing the image matching task without using descriptors. LoFTR is based on SuperGlue’s architecture and thus it employs alternating self-attention and cross-attention layers to process the dense local features extracted by the image via a CNN. The system has four main components. A pair of images is fed to a CNN for extracting coarse-level and fine-level feature maps. Coarse features maps correspond to the $1/8$ or the original image dimension and fine feature maps to the $1/2$. Feature maps are then flattened and with their positional encodings are forwarded to a coarse-level local feature transformer which is composed of alternating self-attention and cross-attention layers. Afterwards, a differential matching layer is utilized for the generation of a confidence matrix which yields a coarse-level match prediction. Finally, coarse matches are refined to generate the final fine-level predictions.

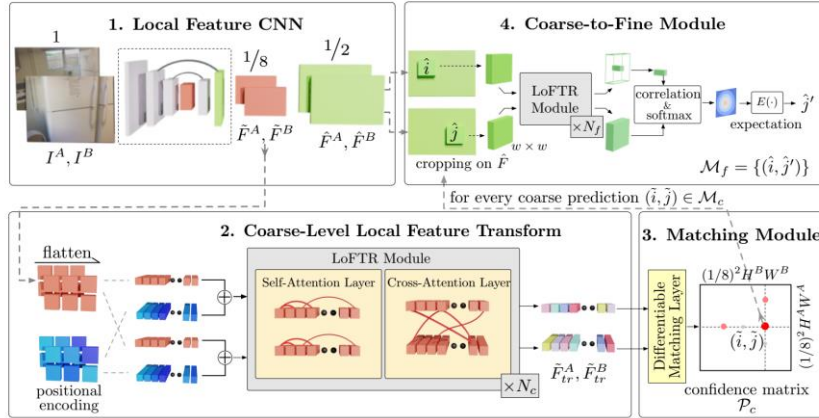


Figure 37: Architecture of solution proposed by LoFTR

Depth Estimation

⁹⁰⁸ J. Sun, Z. Shen, Y. Wang, H. Bao and X. Zhou, "LoFTR: Detector-Free Local Feature Matching with Transformers," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021, pp. 8918-8927, doi: 10.1109/CVPR46437.2021.00881.

Usually, direct methods require depth information for estimating the pose of the ego vehicle. Depth estimation is a very popular area of research regarding pose estimation and deep learning. The proposed algorithms adopt either supervised or unsupervised learning approaches.

The depth estimation using supervised methods was first proposed in the work of Eigen et. al.⁹⁰⁹ where they presented a multi-scale neural network for estimating the depth from a single image. The architecture is split into two halves, one for coarse global depth estimations and one for local refinement. This approach, thus the utilization of the multi-scale features of the image, has been adopted by many works since^{910,911,912}. Depth estimation is usually treated as a regression problem and therefore Liu et al.⁹¹³ proposed a CNN based framework which transforms the depth estimation problem in learning problems. Additionally, depth estimation can be dealt as a classification problem. This is based on the premise that estimating a depth range is simpler than estimating a precise value. Using CNN alone is insufficient for modeling long-range context effectively. Therefore, Grigorev et al.⁹¹⁴ proposed a hybrid network that combines convolutional and ReNet layers. The ReNet layer includes Long Short-Term Memory units, enabling it to capture global context features. Another architecture based on LSTM was proposed to investigate whether RCNN can accurately estimate spatial-temporal depth without geometric consistency or pose supervision⁹¹⁵. Researchers focused on more complex and specific issues to achieve more precise and reliable results⁹¹⁶. They used invariant surface normal to address depth ambiguity⁹¹⁷. In addition, researchers used both sparse depth samples obtained from a low-resolution depth sensor or computed through SLAM algorithms and input images to estimate full-resolution depth. Moreover, a progressive hard-mining network was proposed to overcome the challenge of a large semantic gap and accumulated error over scales⁹¹⁸. DFineNet⁹¹⁹ was proposed to refine depth estimation under sparse and noisy conditions.

The lack of large datasets that provide ground truth labels is the motive behind the researchers that investigated unsupervised solutions on the problem deep learning based depth estimation. A pioneer in this field is the work presented by Garg et al.⁹²⁰ which estimates the depth from a single image. The operation of the network mimics an

⁹⁰⁹ D. Eigen, C. Puhrsch, and R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, vol. 27, 2014

⁹¹⁰ S. Bazrafkan, H. Javidnia, J. Lemley, and P. Corcoran, "Semiparallel deep neural network hybrid architecture: first application on depth from monocular camera," *Journal of Electronic Imaging*, vol. 27, no. 4, 2018

⁹¹¹ Y. Kim, H. Jung, D. Min, and K. Sohn, "Deep monocular depth estimation via integration of global and local predictions," *Ieee Transactions on Image Processing*, vol. 27, no. 8, pp. 4131–4144, 2018.

⁹¹² Z. Y. Zhang, C. Y. Xu, J. Yang, Y. Tai, and L. Chen, "Deep hierarchical guidance and regularization learning for end-to-end depth estimation," *Pattern Recognition*, vol. 83, pp. 430–442, 2018.

⁹¹³ F. Y. Liu, C. H. Shen, G. S. Lin, and I. Reid, "Learning depth from single monocular images using deep convolutional neural fields," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 10, pp. 2024–2039, 2016

⁹¹⁴ A. Grigorev, F. Jiang, S. Rho, W. J. Sori, S. H. Liu, and S. Sai, "Depth estimation from single monocular images using deep hybrid network," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 18 585–18 604, 2017

⁹¹⁵ A. C. S. Kumar, S. M. Bhandarkar, and M. Prasad, "Depthnet: A recurrent neural network architecture for monocular depth prediction," *Proceedings 2018 Ieee/Cvf Conference on Computer Vision and Pattern Recognition Workshops (Cvprw)*, pp. 396–404, 2018.

⁹¹⁶ F. C. Ma and S. Karaman, "Sparse-to-dense: Depth prediction from sparse depth samples and a single image," *2018 Ieee International Conference on Robotics and Automation (Icra)*, pp. 4796–4803, 2018

⁹¹⁷ H. Yan, S. L. Zhang, Y. Zhang, and L. Zhang, "Monocular depth estimation with guidance of surface normal map," *Neurocomputing*, vol. 280, pp. 86–100, 2018.

⁹¹⁸ Z. Zhang, C. Xu, J. Yang, J. Gao, and Z. Cui, "Progressive hard-mining network for monocular depth estimation," *IEEE Trans Image Process*, vol. 27, no. 8, pp. 3691–3702, 2018.

⁹¹⁹ M. I. D. Zhang Y, Nguyen T, "Dfinenet: Ego-motion estimation and depth refinement from sparse, noisy depth input with rgb guidance," *arXiv preprint arXiv:1903.06397*, 2019.

⁹²⁰ R. Garg, B. G. VijayKumar, G. Carneiro, and I. Reid, "Unsupervised cnn for single view depth estimation: Geometry to the rescue," *Computer Vision - Eccv 2016, Pt Viii*, vol. 9912, pp. 740–756, 2016.

autoencoder, in which the encoding part estimates the inverse depth map using the left image and the decoding part reconstructs the original left image using the inverse depth map and the right image. In a similar way, Godard et al.⁹²¹, in a similar way, proposed a way to infer disparities between a pair of left and right images, and thereby acquiring better depth estimation by enforcing depth consistency constraints.

LiDAR odometry

Generally, in the case of deep learning-based LiDAR odometry the point cloud is firstly preprocessed using a spherical projection which results in a multi-channel image. LO-Net⁹²² generates this multi-channel image and subsequently performs end-to-end pose estimation. The information extracted by the image is used for pose regression and for filtering out dynamic objects. Extracted features from the projected point cloud for the relative pose estimation are used in the case of DeepVCP⁹²³. The model proposed in this work is trained from data generated by a Velodyne HDL-64 LiDAR sensor. PWClo-Net⁹²⁴ uses hierarchical embedding mask optimization for refining the estimated pose in a coarse-to-fine approach. An attentive cost is used to determine the motion patterns between two consecutive point clouds and a trainable embedded mask is used to assign weights to each point in the point cloud based on its local motion pattern.

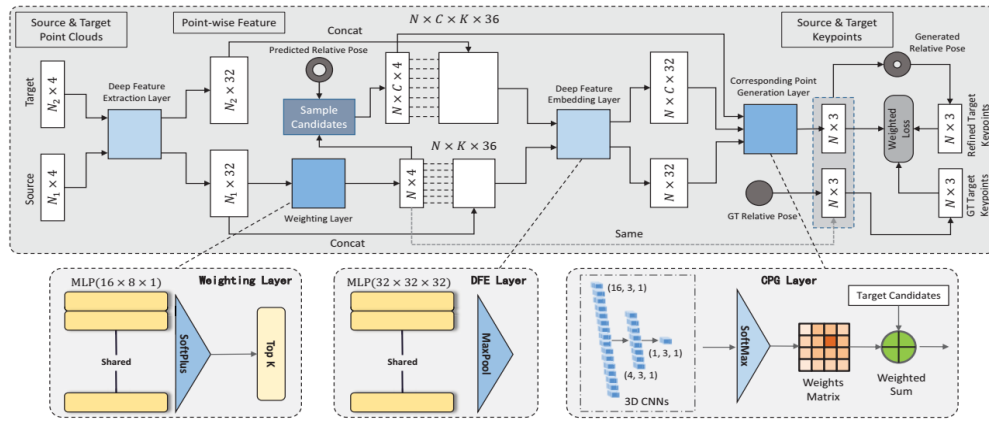


Figure 38: Architecture of DeepVCP

Unsupervised learning can also be employed for this cause, but even though the amount of data needed is significantly less, the proposed algorithms are not as efficient as the ones implemented using supervised learning techniques. For example, Yoon et al.⁹²⁵ proposed an unsupervised learning algorithm which uses a Gaussian

⁹²¹ C. Godard, O. Mac Aodha, and G. J. Brostow, "Unsupervised monocular depth estimation with left-right consistency," 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), pp. 6602–6611, 2017

⁹²² Li, Q.; Chen, S.; Wang, C.; Li, X.; Wen, C.; Cheng, M.; Li, J. Lo-net: Deep real-time lidar odometry. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 8473–8482.

⁹²³ Lu, W.; Wan, G.; Zhou, Y.; Fu, X.; Yuan, P.; Song, S. Deepvcp: An end-to-end deep neural network for point cloud registration. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 12–21

⁹²⁴ G. Wang, X. Wu, Z. Liu and H. Wang, "PWClo-Net: Deep LiDAR Odometry in 3D Point Clouds Using Hierarchical Embedding Mask Optimization," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021, pp. 15905–15914, doi: 10.1109/CVPR46437.2021.01565.

⁹²⁵ Yoon, D.J.; Zhang, H.; Gridseth, M.; Thomas, H.; Barfoot, T.D. Unsupervised Learning of Lidar Features for Use in a Probabilistic Trajectory Estimator. *IEEE Robot. Autom. Lett.* **2021**, *6*, 2130–2138

variational inference setting. The authors claim that their method achieves results comparable to ICP-based methods on the KITTI odometry dataset.

Another approach to DL based LiDAR is the one that combines both traditional and data-driven models. For example, LodoNet⁹²⁶ extracts SIFT features on the projected image and reshapes feature point matchings into vectors. After that, convolutional networks are used for the estimation of the relative poses between the consecutive point clouds. On the other hand, other works^{927,928} follow a different path. More specifically they extract, and match features points on the spherical projection using deep learning-based techniques and then apply ICP.

It must be noted that in all solutions the extraction of robust and repeatable features from low-cost LiDAR sensors (e.g. 16 channel LiDAR) is very challenging due to the sparsity of the point clouds.

Sensor Fusion

Regarding their principle of operation, sensor can be categorized into two groups, proprioceptive sensors that capture the dynamical state and the internal parameters of the dynamic system (e.g. GNSS, gyroscope, imu etc.) and exteroceptive sensors that capture the external environmental conditions (e.g. camera, LiDAR, radar, etc.). Furthermore, sensors can also be categorized depending on whether they're passive, meaning that they capture the surrounding energy, or active when they transmit energy and capture the reflected energy to produce an estimation for their environment.

CNNs fusion systems have gained popularity in the research community. For example, Shi et al.⁹²⁹ described a useful technique to extract the LIDAR projection line from the image and enhanced the visual SLAM-based LIDAR scanning system. Prior to using the extracted image features for pose estimation in visual SLAM, the discovered object was first given an adaptive threshold. Last but not least, a precise and realistic 3D model that can produce 3D point cloud maps with actual colors and true scales was created using the CNN's semantic segmentation method. The work presented by Park et al.⁹³⁰ addresses the problem of the map creation using as input data from uncalibrated LiDAR sensors and stereo camera. It uses two CNNs one for the calibration and one of the deep fusion.

DEEP UNROLLING TECHNIQUES IN SCENE ANALYSIS AND SLAM

Although the literature regarding the deep unrolling paradigm is rich in 2D image and signal processing domain, the benefits of this procedure have not been studied in the automotive domain. In more detail, in the autonomous driving domain, various inverse problems may arise that can significantly affect the performance of scene analysis and SLAM applications. In this project, the goal is to investigate such inverse problems for visual and lidar sensors, designing novel optimization problems and deep unrolling networks with state-of-the-art performance, low-computational complexity, and well-justified-explainable architectures.

⁹²⁶ Zheng, C.; Lyu, Y.; Li, M.; Zhang, Z. Lodonet: A deep neural network with 2d keypoint matching for 3d lidar odometry estimation. In Proceedings of the 28th ACM International Conference on Multimedia, Seattle, WA, USA, 12–16 October 2020; pp. 2391–2399

⁹²⁷ Li, Z.; Wang, N. Dmlo: Deep matching lidar odometry. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October–24 January 2021; pp. 6010–6017

⁹²⁸ Streiff, D.; Bernreiter, L.; Tschopp, F.; Fehr, M.; Siegwart, R. 3D3L: Deep Learned 3D Keypoint Detection and Description for LiDARs. *arXiv* **2021**, arXiv:2103.13808.

⁹²⁹ Shi, Z.; Lyu, Q.; Zhang, S.; Qi, L.; Fan, H.; Dong, J. A Visual-SLAM based Line Laser Scanning System using Semantically Segmented Images. In Proceedings of the 2020 11th International Conference on Awareness Science and Technology (iCAST), Qingdao, China, 7–9 December 2020; pp. 1–6

⁹³⁰ Park, K.; Kim, S.; Sohn, K. High-Precision Depth Estimation Using Uncalibrated LiDAR and Stereo Fusion. *IEEE Transactions Intell. Transp. Syst.* **2020**, *21*, 321–335.

For instance, an interesting problem in the autonomous driving domain that we aim to explore concerns the lidar sensors. In general, lidar sensors play an important role for various navigation problems including localization, object detection and mapping⁹³¹. Different from the typical visible cameras, lidar sensors are able to function properly at night providing visibility at long-range distances. A 3D Lidar sensor utilizing multiple lasers or channels is able to offer dense point clouds, thus capturing fine details of the environment. Nonetheless, the cost of such high-resolution sensors is remarkably high, which restricts the large-scale use of such systems. For instance, the cost of a 64-channel HDL-64E lidar sensor is approximately \$85000. An alternative is to employ a lidar sensor with limited channels e.g., 16. However, this strategy imposes some restrictions, since the 16-channel may provide sparse point clouds without the required detailed information about the environment, thus heavily affecting the performance of lidar odometry⁹³² and object detection⁹³³ tasks.

To this end, considering the high cost of a high-resolution lidar sensor and the limitations and restrictions of low-resolution lidar sensor, in this project, we aim to examine super-resolution methods based on the deep unrolling pipeline, to reconstruct a high-resolution lidar utilizing only as input a low-resolution sensor. Thus, having only a low-resolution lidar with 16 channels instead of using sparse point clouds, we can produce more dense point clouds based on the super-resolution methodologies. Then, the dense point cloud can be used as input to other algorithms for SLAM and scene analysis to increase their performance. Additionally, other inverse problems such as denoising may play a significant part in scene analysis and SLAM tasks as a post-processing procedure.

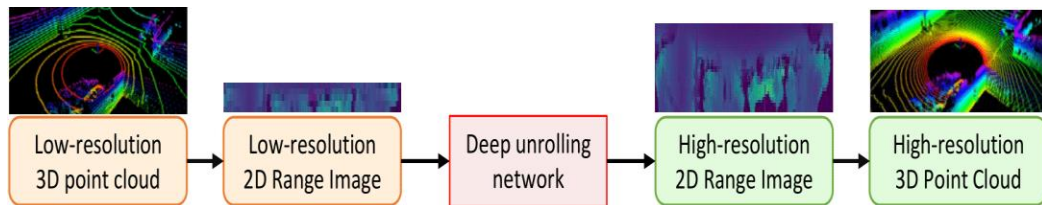


Figure 39: The workflow of a super-resolution lidar methodology. From a sparse 3D point cloud derived from a 16-channel cheap lidar sensor, we can reconstruct a denser point cloud similar to that produced from a 64-channel lidar based on the deep unrolling method

FL-BASED INTELLIGENCE FOR CONNECTED AND AUTONOMOUS VEHICLES

Literature review of FL and automotive domain

Although the FL framework is well-studied in various disciplines e.g., image processing, computer vision, in the autonomous driving domain it is highly under-explored. Considering the distributed nature of the FL along with its privacy preserving properties, the FL pipeline can be utilized in various autonomous driving applications ranging from scene analysis and object detection to SLAM and other inverse imaging problems such as lidar-super-resolution.

⁹³¹ .Shan, J. Wang, F. Chen, P. Szenher, and B. Englot, "Simulation-based lidar super-resolution for ground vehicles," *Robotics and Autonomous Systems*, vol. 134, p. 103647, dec 2020.

⁹³² . Shan and B. Englot, "Lego-loam: Lightweight and ground- optimized lidar odometry and mapping on variable terrain," in

2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018, pp. 4758–4765.

⁹³³ Y. Wu, Y. Wang, S. Zhang and H. Ogai, "Deep 3D Object Detection Networks Using LiDAR Data: A Review," in *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1152-1171, 15 Jan.15, 2021, doi: 10.1109/JSEN.2020.3020626.

In literature only several studies have investigated the benefits derived from the FL framework. Jallepalli et. al⁹³⁴. utilized the baseline FedAvg algorithm to tackle the problem of object detection in the open-source KITTI Vision Benchmark dataset. Each client employed a YOLOV3 object detection model to train based on its local dataset. After the local training, the clients broadcast their model to the centralized server, which performs a simple weighted average aggregation rule, thus deriving the global model. The above study concluded that the FL framework is able not only to obtain similar results with the classical centralized deep learning but also to speed up the local models training process. Similar finding derived from study⁹³⁵, which employed the FedAvg approach aiming to design accurate local models to predict the wheel steering angle in self-driving vehicles. Moreover, study⁹³⁶ also examined the impact of FL on the above-mentioned problem. However, different from study⁷⁰, this method proposed a decentralized FL scheme, thus removing the need for a central server allowing peer-to-peer communication among the clients.

Moving on, study⁹³⁷ provided a more theoretical and practical analysis of the FL on vehicular networks, providing a discussion regarding the training procedure of the local models such as the data distribution of the clients' data and the non i.i.d. nature of the autonomous driving datasets. Additionally, the authors highlight the communication aspects of the FL framework i.e., data rate, reliability, transmission overhead, privacy and resource management. Study⁹³⁸ considered a semantic segmentation task under the FL framework. In this study, a benchmark platform was developed including two datasets and several state-of-the-art FL algorithms. The goal of this study was to examine the statistical heterogeneity of the autonomous driving datasets. In more detail, drivers from different countries and areas may experience different weather conditions or landscapes e.g., cityscapes instead of rural areas. Thus, similar challenges imposed by the non i.i.d. nature of the data in other FL applications can also occur in autonomous driving. In view of this, FL architectures suitable to handle the statistical heterogeneity of the local clients are imperative to developed for realistic autonomous driving applications, such as the method developed in study⁹³⁹, where a cluster-based FL algorithm was proposed.

Beyond SoA: cooperative perception

Conventionally, DNN training is based on centralized learning, which collects the datasets from autonomous vehicles, trains the DNNs at the cloud, and deploys trained DNNs on vehicles for inference. This paradigm is effective in closed driving areas where the owner of data collection vehicles is also the DNN provider. However, future perception systems must cope with multi-variate open scenarios, which involve corner cases due to infinite

⁹³⁴ D. Jallepalli, N. C. Ravikumar, P. V. Badarinath, S. Uchil and M. A. Suresh, "Federated Learning for Object Detection in Autonomous Vehicles," 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, United Kingdom, 2021, pp. 107-114, doi: 10.1109/BigDataService52369.2021.00018.

⁹³⁵ H. Zhang, J. Bosch and H. H. Olsson, "End-to-End Federated Learning for Autonomous Driving Vehicles," 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 2021, pp. 1-8, doi: 10.1109/IJCNN52387.2021.9533808.

⁹³⁶ A. Nguyen et al., "Deep Federated Learning for Autonomous Driving," 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 2022, pp. 1824-1830, doi: 10.1109/IV51971.2022.9827020.

⁹³⁷ A. M. Elbir, B. Soner, S. Çöleri, D. Gündüz and M. Bennis, "Federated Learning in Vehicular Networks," 2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Athens, Greece, 2022, pp. 72-77, doi: 10.1109/MeditCom55741.2022.9928621.

⁹³⁸ L. Fantauzzo et al., "FedDrive: Generalizing Federated Learning to Semantic Segmentation in Autonomous Driving," 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 2022, pp. 11504-11511, doi: 10.1109/IROS47612.2022.9981098.

⁹³⁹ A. Taïk, Z. Mlika and S. Cherkaoui, "Clustered Vehicular Federated Learning: Process and Optimization," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 12, pp. 25371-25383, Dec. 2022, doi: 10.1109/TITS.2022.3149860.

scenario space and visual occlusions⁹⁴⁰. To address these challenges a lifelong multi-stage training that updates the DNN parameters whenever rare or occluded objects are detected can be an attractive alternative. In this case, sensor data is inherently distributed at customers' vehicles and contains high-resolution human-related private information, leading to the potential of sensitive information leakage⁹⁴¹. Federated (deep) learning assisted autonomous vehicles (FLAV) is an emerging paradigm to overcome the privacy issue by training DNNs via parameter and output aggregation instead of dataset aggregation⁹⁴²: parameter aggregation leverages vehicular networks to migrate knowledge among different vehicles⁹⁴³; output aggregation leverages road sensors' perception results to annotate occluded objects for local parameter updates⁹⁴⁴. The performance of FLAV highly depends on the associated network resources and road sensor placements. Existing works on FLAV design these factors from either a driving or networking perspective, which ignores the interdependency between driving tasks (i.e., Consumer) and vehicular networks (i.e., data provider). This research gap has been identified as the core issue in autonomous vehicles systems⁹⁴⁵. Therefore, TRUSTEE envisions to integrate driving and networking features for system-level FLAV perception under practical network resource constraints. Frameworks of vehicle-edge-cloud networking, multi-stage DNN training, and multi-task generation, will also be taken into account for the design of the proposed FL-based perception system.

Beyond SoA: cooperative SLAM

Even if we are far away from a deep-based end-to-end SLAM system, multiple works present novel and innovative ideas which use deep learning techniques to substitute components of traditional SLAM systems like the feature extraction and matching, and the depth estimation. However, these techniques usually require a huge amount of annotated data for training purposes. Decentralized solutions based on the FL scheme provide solutions which overcome privacy issues and concerns. We plan to investigate cooperative SLAM solutions by deploying federated deep learning-based algorithms for feature extraction and matching algorithms without transferring raw data among the agents. Finally, we are going to evaluate the performance of the models trained in a federated way against similar models following a centralized scheme.

Beyond SoA: FL and deep unrolling for inverse problems in automotive domain

As we mentioned above, in the autonomous driving domain, various inverse problems may arise that can significantly affect the performance of scene analysis and SLAM applications. Based on the deep unrolling framework, we can provide a novel bridge between the optimization-based methods and the deep learning approaches to design deep architectures, which combine the advantages of both worlds. In other words, we can depart from the ad-hoc, intuition-based design of deep learning models and move towards well-justified architectures derived from the prior domain knowledge of the examined problem. These models are highly interpretable since their architectures (e.g., number of layers) and learnable parameters derived from the solution of

⁹⁴⁰ D. Feng et al., "Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1341–1360, Mar. 2021.

⁹⁴¹ S. Savazzi, M. Nicoli, M. Bennis, S. Kianoush, and L. Barbieri, "Opportunities of federated learning in connected, cooperative and automated industrial systems," *IEEE Commun. Mag.*, vol. 59, no. 2, pp. 16–21, Feb. 2021.

⁹⁴² B. Liu, L. Wang, M. Liu, and C. Xu, "Federated imitation learning: A novel framework for cloud robotic systems with heterogeneous sensor data," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 3509–3516, Apr. 2020.

⁹⁴³ J. Posner, L. Tseng, M. Aloqaily, and Y. Jararweh, "Federated learning in vehicular networks: Opportunities and solutions," *IEEE Netw.*, vol. 35, no. 2, pp. 152–159, Mar./Apr. 2021.

⁹⁴⁴ E. Arnold, M. Dianati, R. de Temple, and S. Fallah, "Cooperative perception for 3D object detection in driving scenarios using infrastructure sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 1852–1864, Mar. 2022.

⁹⁴⁵ I. Yaqoob, L. U. Khan, S. M. A. Kazmi, M. Imran, N. Guizani, and C. S. Hong, "Autonomous driving cars in smart cities: Recent advances, requirements, and challenges," *IEEE Netw.*, vol. 34, no. 1, pp. 174–181, Jan./Feb. 2020.

an optimization algorithm. As a future direction our plan is to incorporate the deep-unrolling networks into novel FL methodologies utilizing the distributed nature and privacy preserving properties of the FL pipeline. Considering that different clients such as autonomous vehicles contain different datasets, these clients can collaborate in a privacy preserving manner to obtain a global model, which captures valuable information across the clients. The derived global model can be employed to solve inverse problems in the automotive domain such as denoising and lidar super-resolution, which are crucial preprocessing steps for SLAM and scene analysis tasks. A great benefit of the connection of deep unrolling with the FL framework is the fact that the proposed deep unrolling networks due to their concise nature have low computational complexity. Thus, the communication cost between the server and the clients are notably smaller, since the models contain significantly less learnable parameters as compared to other deep learning architectures.

Beyond SoA: automotive dataset for federation learning

In the context of this project, we are going to provide a fully integrated simulator stack that will be capable of providing synchronized and annotated data, mainly targeting at FL techniques for automotive scenarios. The nature and the flexibility of the already implemented integrated simulator gives us the opportunity to provide a dataset that can have the following properties:

- It should contain data from multiple sources like raw sensor data (e.g. images, point clouds, depth images), communication messages (V2X messages compliant with the ETSI ITS-G5 protocols), and the results of running algorithms (e.g. vehicle's estimated path, 2D detections and bounding boxes, 3D detections and bounding boxes, extracted features points etc.).
- The dataset should be fully annotated.
- Since the integrated simulator gives us the capability to execute different automotive scenarios, the data should have a diversity and must be representative of scenarios and cases that can be encountered in real world driving.
- The dataset can provide a realistic data distribution concerning for example geographic locations, weather conditions, and driving environments.
- The data should be balanced to avoid the bias in training. This means that for example in the case of detection algorithms it should contain an equal number of classes.

The size of the dataset should be large enough to ensure the effective training of the machine model. The size should be depended on the machine learning algorithms and the complexity of the problem under investigation.

SOTA ON TRUSTEE-RELATED REFERENCE ARCHITECTURES

The creation of system, solution, and application architectures is guided by a reference architecture. It offers uniform definitions for the system of interest, its breakdowns and design patterns, as well as a common vocabulary for discussing implementation specification and contrasting options. A reference architecture, in other words, offers a standard framework around which more in-depth discussions can be centered. It allows for the identification and understanding of the most significant problems and patterns across its applications in numerous different use cases by remaining at a higher level of abstraction. Reference architectures avoid specifics, which also makes it possible for subsequent designs to adhere to them without being constrained by arbitrary restrictions.

This section contains an overview of relevant state of the art reference architectures and reference architecture models related to the scope and planned results of the TRUSTEE project.

REFERENCE ARCHITECTURE MODEL, REFERENCE ARCHITECTURE, CONCRETE ARCHITECTURE, AND SYSTEM DEFINITIONS

According to IEEE/ISO/IEC 42010-2011⁹⁴⁶, architecture represents “fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and in the principles of its design and evolution”. The term system (as described in ISO/IEC15288) refers to entities whose architectures are of interest, that are “man-made and may be configured with one or more of the following: hardware, software, data, humans, processes, procedures, facilities, materials, and naturally occurring entities”.

A reference model is described as “an abstract framework for understanding significant relationships among the entities of some environment” by OASIS (Organization for the Advancement of Structured Information Standards)⁹⁴⁷. The Reference Architecture Model (RAM) is an abstract framework made up of related concepts that promotes common understanding and logical groupings within architectures.

According to Bass et al., Reference Architecture (RA) is a “reference model mapped onto software elements (that cooperatively implement the functionality defined in the reference model) and the data flows between them”. Reference architectures are designed to capture the key characteristics of systems’ (concrete) architectures in a particular domain. However, it is frequently unclear what they should contain and which fundamental concerns they should address.

Architecture descriptions should identify:

- System stakeholders (including users, operators, owners, developers, maintainers);
- Fundamental concerns (including the purpose of the system, suitability of the architecture to fulfill the set objective, feasibility, risks, maintainability, evolution);
- Architecture views (representing a related set of concerns as seen from a perspective a view is taken, a viewpoint); and
- The rationale for each important architecture decision.

The architecture descriptions, however, are frequently fuzzy in reality. A reference architecture model, or RAModel (Reference Architecture Model), was proposed by Nakagawa et al. in an effort to better understand RA. They came to the conclusion that an RA for a particular application domain should address business rules, architectural styles (addressing quality attributes), best practices for software development, and technical, business, and customer context knowledge. Furthermore, they classified the elements of RA based on a domain (legislation, quality, compliance, regulation), application (goals, needs, risk, scope, functional requirements, domain data), infrastructure (software and hardware elements, best practices), and cross-cutting (internal and external communication, domain terminology, decisions with rationale).

Concrete Architecture (CA) can be considered an instance of RA for a specific context and addresses particular requirements for involved stakeholders. It is specific to the implementation, defining what the system will do and how. In order to define boundaries for implementing a concrete system, the definition of CA involves making numerous different design decisions.

The relation between RAM, RM, CA, and the system, i.e., the implementation of the CA, is shown in Figure 40.

⁹⁴⁶ ISO; IEC; IEEE. Systems and software engineering—Architecture description. ISO/IEC/IEEE 2012, 42010, 1–46.

⁹⁴⁷ Reference Model for Service Oriented Architecture 1.0. OASIS Stand. 2006, 12, 18

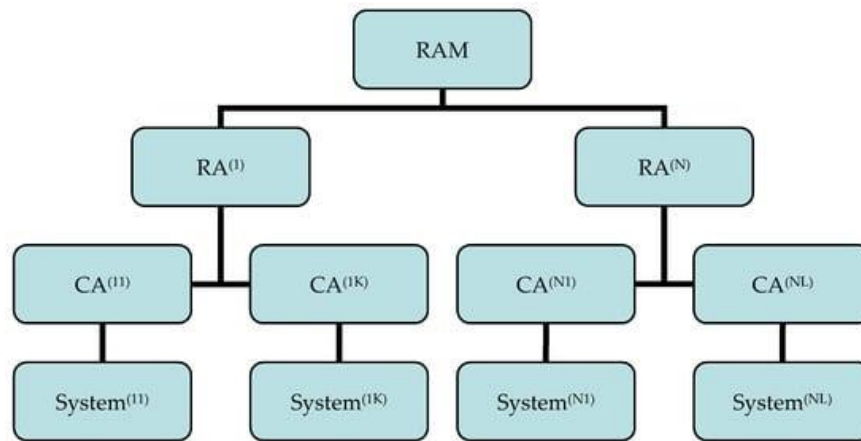


Figure 40: Relation between RAM, RA, CA, and the System [image from⁹⁴⁸]

RELEVANT STANDARDS

ISO/IEC/IEEE 42010

When considering complex system architectures, multiple stakeholders are involved. These stakeholders have many intertwined interests related to the system and their concerns cover its entire lifecycle. The complexity of the system requires a framework to identify and classify stakeholder concerns into appropriate categories. Such a framework enables a systematic evaluation of systems, as well as the resolution required to design and build them.

The Industrial Internet Consortium (IIC) defined the Industrial Internet Architecture Framework (IIAF) using "ISO/IEC/IEEE 42010:2011" to meet this need. The IIAF defines conventions, principles and practices for a consistent description of IIoT architectures. This standard-based architecture framework facilitates systematic and effective evaluation and resolution of stakeholder concerns. Hence, this architectural framework serves as a valuable resource for Industrial Internet Reference Architecture development guidance, documentation, and communication.

ISO/IEC/IEEE 42010 was prepared by Joint Technical Committee ISO/IEC JTC 1, Information technology, Subcommittee SC 7, Software and systems engineering, in cooperation with the Software and Systems Engineering Standards Committee of the IEEE Computer Society, under the Partner Standards Development Organization cooperation agreement between ISO and IEEE. It is an international standard for architectural descriptions of systems and software and provides a blueprint for data abstraction and protection, security, privacy, and safety. ISO/IEC/IEEE 42010:2011 focuses on creating, analyzing, and sustaining system architecture through the use of architecture descriptions. The standard establishes a conceptual model of architecture description and specifies the required contents of an architecture description. Architecture viewpoints, architecture frameworks and architecture description languages are introduced for codifying conventions and common practices of architecture description. The required content of architecture viewpoints, architecture frameworks and architecture description languages is specified. Figure 41 presents the architecture description defined by the standard, used to describe the architecture of a system of interest.

⁹⁴⁸ Grguric, A.; Khan, O.; Ortega-Gil, A.; Markakis, E.K.; Pozdniakov, K.; Kloukinas, C.; Medrano-Gil, A.M.; Gaeta, E.; Fico, G.; Koloutsou, K. Reference Architectures, Platforms, and Pilots for European Smart and Healthy Living—Analysis and Comparison. *Electronics* **2021**, *10*, 1616. <https://doi.org/10.3390/electronics10141616>

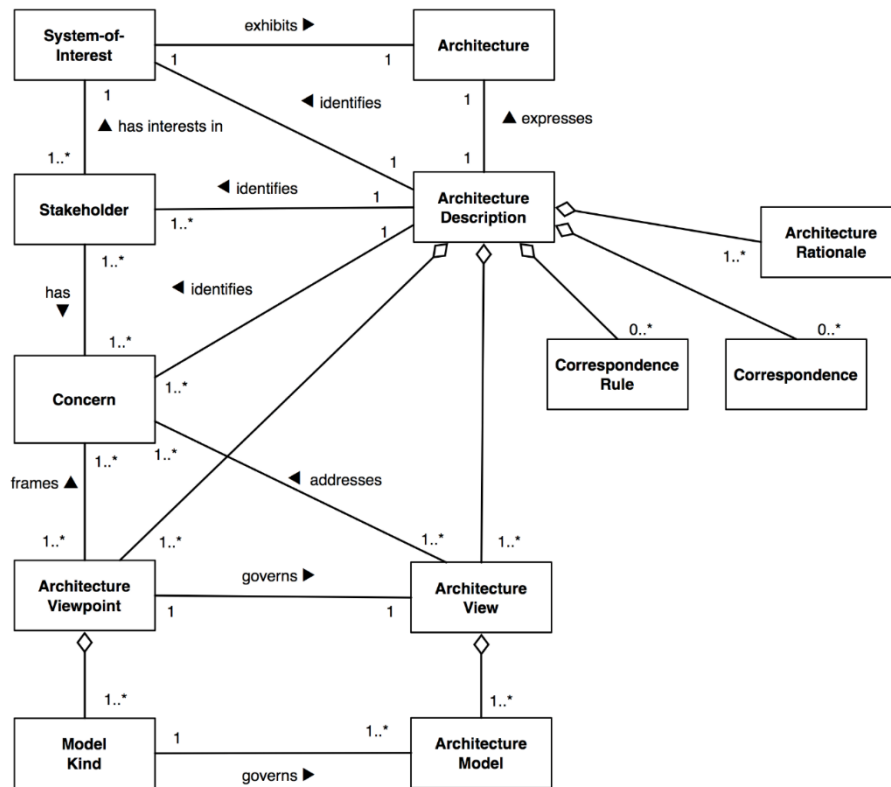


Figure 41: ISO/IEC/IEEE 42010:2011—Architecture Description [image from v1.9 IIRA⁹⁴⁹]

The second edition of the standard, ISO/IEC/IEEE 42010:2022, was published in November 2022, as the previous version ISO/IEC/IEEE 42010:2011 has been revised.

ISO/IEC 30141:2018

A standardized IoT Reference Architecture using industry best practices, a common vocabulary, and reusable designs is provided by ISO/IEC 30141:2018. It uses a top-down approach, starting with collecting the most important characteristics of IoT, extracting them into a generic IoT Conceptual Model, and creating a high-level system-based reference model with five architecture views from different perspectives: functional view, system deployment view, networking view and usage view. Figure 42 illustrates the relations between the stated architecture views, Reference Model and Domain concept.

⁹⁴⁹ The Industrial Internet of Things Volume G1: Reference Architecture, June 19, 2019, available online <https://www.iiconsortium.org/pdf/IIRA-v1.9.pdf> [Accessed 7.11.2022]

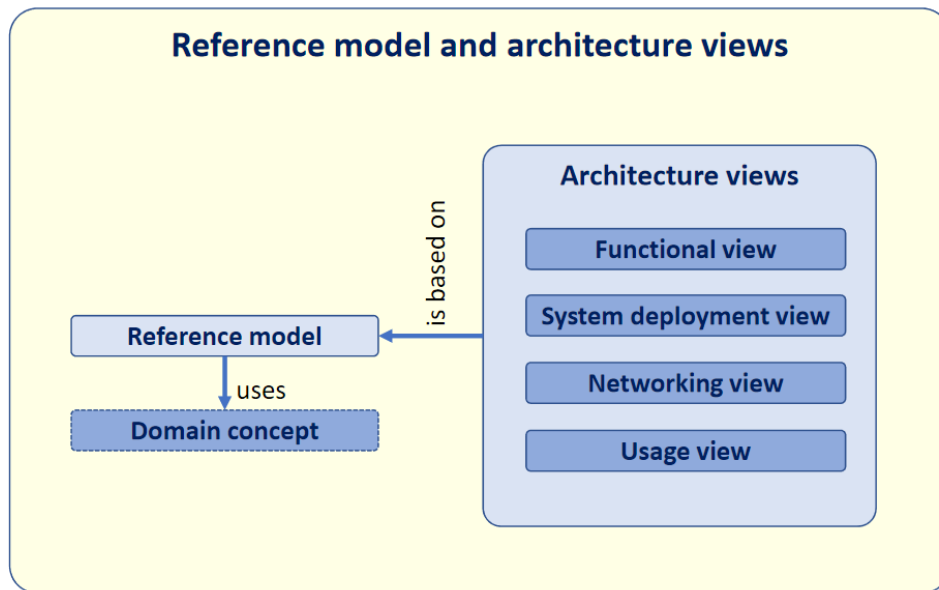


Figure 42: ISO/IEC 30141 Reference model and views

The functional view, which is independent of technology and focuses on the concepts of intra and cross domain functions, describes the interdependencies among the functions that make up a system and supports the activities described in the user view.

The system view describes the generic components, such as the networks, devices, and subsystems needed to build an IoT system, and describes it using its physical components, such as subsystems, devices, and networks.

The usage view addresses the activities, roles and sub-roles, services, and cross-cutting aspects of the IoT system from a user perspective as well as how it is developed, tested, operated, and used. The information view conceptually defines information for entities as pure values devoid of relevant or useable context.

The main communication networks used by IoT systems and the entities they connect to are described in the networking view. These include user networks, services, proximity, and access.

REFERENCE ARCHITECTURE MODELS, WITH INDUSTRIAL AND CROSS-SECTOR FOCUS

IIRA – INDUSTRIAL INTERNET REFERENCE ARCHITECTURE

The Industrial Internet Consortium (IIC) released the latest version of the open, standards-based Industrial Internet Reference Architecture (IIRA) for IoT systems in June 2019. IoT participants can use the IIRA as a common framework to guide development, documentation, communication, and deployment since it specifies models, definitions, and a well-defined vocabulary.

The IIRA standard identifies the fundamental architecture constructs in form of viewpoints and models and adapts architectural approaches from the ISO/IEC/IEEE 42010-2011 Systems and software engineering - Architecture description standard.

These four defined viewpoints are:

- Business viewpoint
- Usage viewpoint

- Functional viewpoint
- Implementation viewpoint

In terms of vision, values, and objectives in a business and regulatory context, the business viewpoint attempts to address the business stakeholders' concerns.

The usage viewpoint describes the activities that coordinate different units of work over different system components and is related to how a system realizes the key capabilities identified in the business viewpoint. These actions provide information for specifying system requirements.

The functional viewpoint, which addresses issues with the functional capabilities and structure of the system and its components, is arguably the most crucial in the specification. The perspective outlines the breakdown of a typical system into five functional domains while emphasizing the key components. Which are:

- Control domain
- Operations domain
- Information domain
- Application domain
- Business domain

In addition to functional domains, crosscutting functions and system characteristics are essential for describing an IIoT system (Figure 43). Crosscutting (or enabling) functions are defined as additional functions that must be offered across many of the functional components (domains) of the system in order to enable the major system functions. Furthermore, emergent system-wide properties or system characteristics are an inherited component of a complex system since an IIoT system is not simply the sum of what is provided by its constituent functional components, but also includes emergent behaviours or properties brought about by the interactions of the constituent parts.

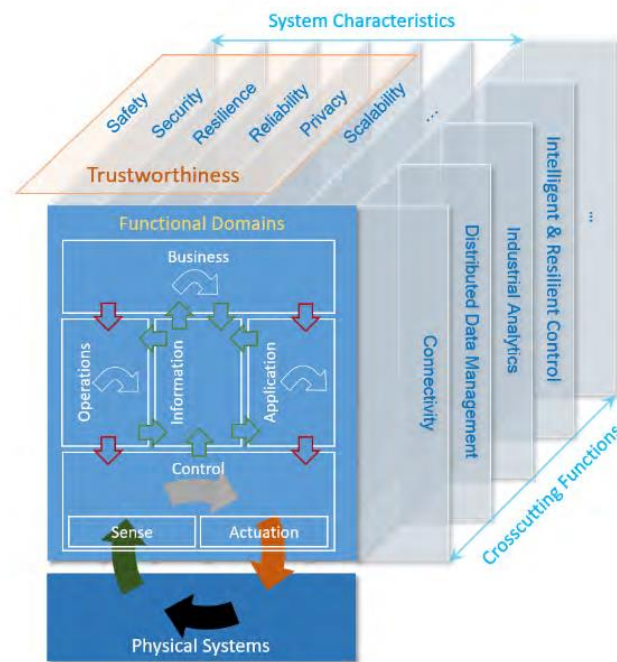


Figure 43: Functional domains, crosscutting functions and system characteristics⁹⁵⁰

The implementation viewpoint outlines the issues with implementing the system's capabilities and structure. It offers a technical explanation of how the system and its parts implement functions and activities that are defined by the usage and functional viewpoints. It also takes into account requirements defined in the business viewpoint. Three example architectural patterns for implementing a system are provided in the IIRA document:

- Three-tier architecture pattern
- The gateway-mediated edge connectivity and management architecture pattern
- Layered databus pattern

The edge, platform, and enterprise tiers make up the three-tier architectural pattern. A local connectivity solution for a system's edge is included in the gateway-mediated edge connectivity and management architecture pattern, which bridges to a wide area network. Low-latency, secure, peer-to-peer data communications across logical layers of the system are provided by the layered databus.

IOT REFERENCE MODEL BY ITU-T Y.2060

The International Telecommunication Union (ITU) is a specialized agency of the United Nations tasked with researching technical, operational, and tariff issues and making recommendations for them in an effort to harmonize telecommunications on a global scale. The work is divided up into various Study Groups (SG), which perform standardization work and create standards recommendations for the various international telecommunications fields. The Focus Groups support the Study Groups by offering a different working environment for the swift

⁹⁵⁰ <https://www.iiconsortium.org/IIRA/>

development of specifications in their chosen fields and address industry needs as they arise, if they are not covered by current Study Groups.

The ITU-T Y.2060 recommendation, which gives an overview of the Internet of things (IoT), was released in 2012 by the SG 13. The ITU-T Y.2060 standard defines the IoT concept and scope, as well as its fundamental traits, high-level requirements, and reference model. The described reference model has two verticals (management and security) in addition to four layers (application, service and application support, network, and device). [10]. Figure 44 shows the ITU-T Y.2060 IoT reference model.

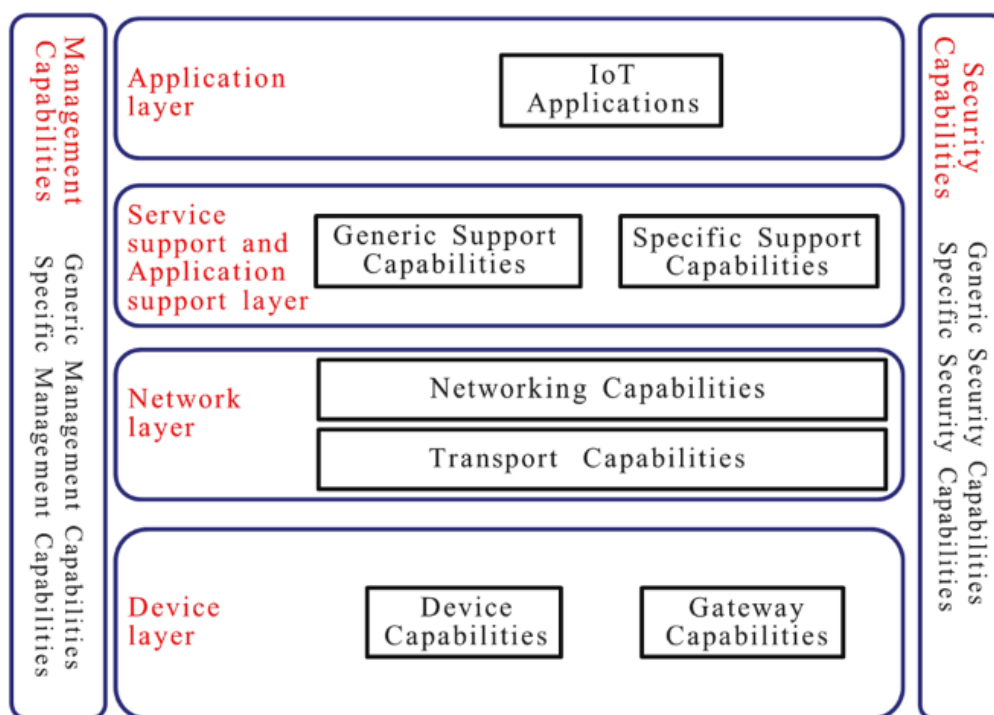


Figure 44: IoT reference model⁹⁵¹

The device layer of ITU-T Y.2060 describes two kinds of capabilities: device and gateway capabilities. While gateway capabilities are related to multiple interface support and protocol conversion, device capabilities are related to gathering and directly or indirectly uploading information.

Networking and transport capabilities are part of the network layer. Networking capabilities are important network connectivity control functions. Information about IoT services and applications is transported using the transport capabilities.

Service support and application support layers are related to generic and specific support capabilities. Generic capabilities are common capabilities that can be used by different IoT applications, whereas specific support capabilities are capabilities that address the requirements of diversified applications.

⁹⁵¹ SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT-GENERATION NETWORKS Next Generation Networks – Frameworks and functional architecture models https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-Y.2060-201206-I!!PDF-E&type=items

IoT applications are found in the application layer.

The two verticals include general fault tolerance, configuration, authentication, and authorization functionalities.

CREATE-IOT 3D RAM

CREATE-IoT 3D RAM, a result of a project CREATE-IoT (CROSS fERTILISATION through ALIGNMENT, synchronization and EXCHANGES for IoT)⁹⁵² with the goal of stimulating collaboration between IoT initiatives and supporting the development of IoT ecosystems based on open technologies and platforms.

The IoT European Large-Scale Pilots Programme (LSP) was launched in 2016 to foster the deployment and evolution of Internet of Things (IoT) solutions through the integration of advanced IoT technologies. It includes the innovation consortia of several funded projects that are applying IoT approaches to specific real-life challenges across different use cases, from autonomous driving, smart city, active and health aging, and farming¹⁰.

Coordinated by the work within CREATE-IoT project, the initiative designed LSP RA model to consolidate all the architectures of the five IoT LSP projects. Apart from CREATE-IoT another supporting project in this work was U4IoT which aimed to actively engage end-users and citizens to achieve IoT societal acceptance. As described in the report, the proposed IoT Reference Architecture is building on the common approach to "horizontal" functional layering that reflects the IoT implementations across various domains, and is expanded with two additional dimensions, cross-cutting functions and properties that are partly addressed by LSPs, while including other views of the IoT systems. This three-dimensional model, presented in Figure 45 is aligned with several standard IoT architecture model, such as RAMI 4.0 and IIRA. Below is a more detailed description of each dimension in LSP RA.

⁹⁵² <https://european-iot-pilots.eu/project/create-iot/>



Figure 45: The general structure of CREATE-IoT 3D RAM ⁹⁵³

The Layers dimension represents common horizontal layering which includes:

- Physical Layer – related to devices such as sensors, actuators and other wired or wireless things
- Network Communication Layer – responsible for providing connectivity physical layer with IoT nodes in a cloud-based platform through gateways and other communication and processing units.
- Processing Layer – related to edge computing, data element analysis and transformations. It enables remote maintenance and management and processing of the data at the edge (as opposed to cloud)
- Storage Layer – responsible for data accumulation using database capabilities
- Abstraction Layer – related to aggregation, access and rules engine to produce higher level data and provides data normalization, reformatting, cleaning and simple statistics
- Service Layer - provides an integrated development environment to simplify the development of applications, analytics and service component for more complex processing
- Application Layer – responsible for presenting data with enhanced visualizations and interactive dashboards
- Collaboration and Process layer - enables integration with existing enterprise and other external systems

⁹⁵³ Vermesan, O.; Eisenhauer, M.; Serrano, M.; Guillemain, P.; Sundmaeker, H.; Tragos, E.Z. The Next Generation Internet of Things–Hyperconnectivity and Embedded Intelligence at the Edge. In Next Generation Internet of Things. Distributed Intelligence at the Edge and Human Machine-to-Machine Cooperation; River Publishers: Gistrup, Denmark, 2019.

The second dimension is Cross-cutting Functions which address additional functionalities that are not linked to a single layer but whose provision requires spanning across several layers. This includes security, privacy, reliability, etc.

The third dimension are Properties that addresses the global properties of the IoT system that are not provided by a proper implementation of functions (at all layers) and cross-cutting functions. This includes availability, scalability interoperability, etc.

RAMI4.0 – REFERENCE ARCHITECTURE MODEL INDUSTRIE

To support Industry 4.0 initiatives whose ideas, structures, and procedures are being adopted globally to modernize manufacturing, the German Electrical and Electronic Manufacturers' Association (ZVEI) created the Reference Architecture Model for Industry 4.0 (RAMI 4.0) in 2015. This means that the development of RAMI 4.0 focused on industrial production as the primary area of application, including discrete manufacturing to process industries.

The RAMI 4.0 is a three-dimensional model describing important aspect of Industry 4.0 in terms of life-cycle value streams, manufacturing hierarch levels and IT system itself (Figure 46). The IT perspective is described with six layers ranging from asset as the lowest layer to business as a top layer. The asset layer represents reality and is related to physical components of the system including a human. This layer represents physical things that will be virtually represented in the remaining layers. The integration layer connects real and virtual representations of the world by generating easy to process content in a form that can be processed by computer. The communication layer is responsible for the transmission of data in a uniform manner but also allows control of the integration layer. The information layer is responsible for processing, integrating and persisting the data and events, as well as for describing the data related to the technical functionality of an asset. It is also responsible for the provision of structured data via interfaces. The functional layer can be seen as platform for horizontal integration of various functions. Rules, decision-making logic, remote access, and horizontal integration are generated inside this layer. Finally, the business layer is responsible for the orchestration of the services provided by the functional layer and for mapping the services to the business model and the business process models.

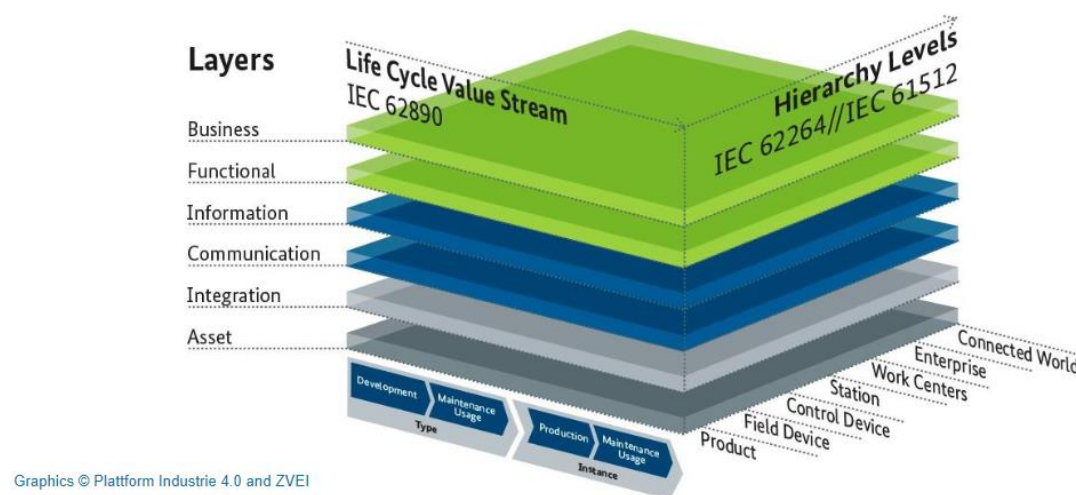


Figure 46: IT perspective layers

The function of each layer is:

- The asset layer describes the physical components of a system, for example, production equipment, product part, sensors, documents, as well as humans. For every asset represented in this layer, there must be a

virtual representation in the above layers. Among the physical assets, this layer includes the digital interface with humans and the relationship to elements in the integration layer.

- The integration layer deals with easy-to-process information content and can be considered as a bridge between the real and the IT world. It contains all elements associated with the IT, including field buses, HMIs, necessary to implement a function, as well as the properties and process-related functions required to use an asset in the intended way and generates events based on the acquired information.
- The communication layer is responsible for the standardized communication between the integration and information layer. Therefore, it performs the transmission of data and files and standardizes the communication from the Integration Layer, providing uniform data formats, protocols and interfaces in the direction of the Information Layer. It also provides services to control the integration layer.
- The information layer holds the necessary data in a structured and integrated form and provides the interfaces to access this structured data from the functional layer. It is responsible for processing, integrating and persisting the data and events, as well as for describing the data related to the technical functionality of an asset. It can be considered the run-time environment for Complex Event Processing (CEP) where rule-based (pre-) processing of events takes place, data APIs and data persistence mechanisms. So, events are received from the communication layer, transformed and forwarded accordingly.
- The functional layer describes the logical and technical functions of an asset providing a digital description of its functions and a platform for the horizontal integration of various functions; it also describes the business model mapping, and business processes which can be adjusted based on inputs from the functional layer, providing models with runtime data of processes, functions and applications.
- The business layer is in charge to orchestrate the services provided by the functional layer. It maps the services to the business (domain) models and the business process models. It also models the business rules, legal and regulatory constraints of the system. The processes to ensure of the economy are located on this level.

The life-cycle value stream dimension is based on IEC 62890, Life-cycle management for systems and products, used in industrial-process measurement, control, and automation where “Type” and “Instance” are distinguished as two main concepts; a “Type” becomes an “Instance” when design and prototyping have been completed and the actual product is being manufactured.

The hierarchy levels dimension, shown on the right horizontal axis, is based on IEC 62264 and IEC 61512 standards, the international standards series for enterprise IT and control systems, but extended with two new levels. At the bottom is the “Product” which is related to elements within the production system, and at the top there is the “Connected World” which represents its outer networks or the ecosystem.

There are clear benefits of RAMI 4.0, especially for the manufacturing industry and the specification can be regarded as a map of Industry 4.0 solutions. However, RAMI4.0 supports different standards in different layers (companies can use unofficial or internal standards) and there are some inconsistencies regarding the concept of the Layers and how the Hierarchy Levels fit the Life Cycle & Value Stream.

REFERENCE ARCHITECTURE MODELS, WITH A FOCUS ON DATA SPACES

IDS RAM – INTERNATIONAL DATA SPACES ASSOCIATION, REFERENCE ARCHITECTURE MODEL 4.0

More than 101 businesses and organizations from 20 different countries, representing a range of sizes and industries, make up the International Data Spaces Association (IDSA). By using an open, vendor-independent architecture for a peer-to-peer network that offers usage control over data from all domains, the IDSA seeks to ensure data

sovereignty. The main goal is to establish a de facto market standard for the sovereign use of data on a global scale. The IDS's activities are closely aligned with Platform Industry 4.0.

The IDS Reference Architecture Model (RAM 3.0) released in April 2019 uses a five-layer structure to express different stakeholder concerns and points of view in accordance with widely used system architecture models and standards (such as ISO 42010, 4+1 view model). These are the layers (Figure 47):

- The Business Layer: specifies and categorizes the different roles and the main activities and interactions between roles
- The Functional Layer: related to the functional requirements and features
- The Process Layer: specifies the interactions taking place between the different components of and it provides a dynamic view
- The Information Layer: defines a conceptual model which makes use of linked-data principles for describing both the static and the dynamic aspects
- The System Layer: concerned with the decomposition of the logical software components, considering aspects such as integration, configuration, deployment, and extensibility

Additionally, the Reference Architecture Model is composed of three perspectives that are implemented across all five layers: Security, Certification, and Governance.

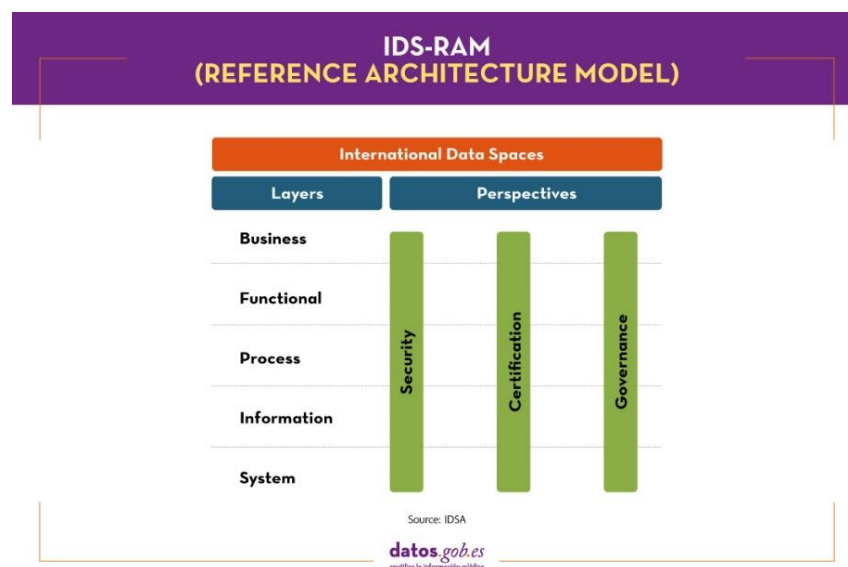


Figure 47: IDS-RAM ⁹⁵⁴

⁹⁵⁴ <https://internationaldataspaces.org/publications/ids-ram/>
https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0

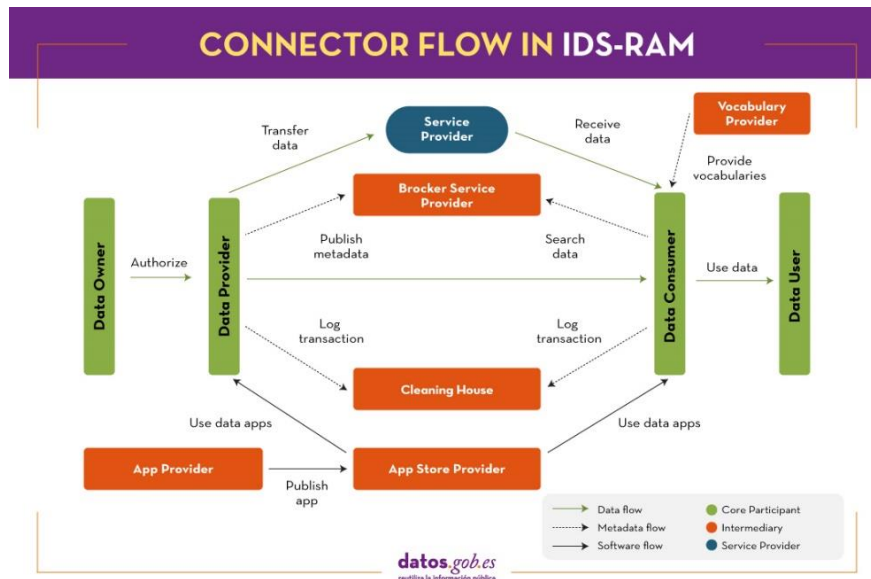


Figure 48: Connector flow in IDS-RAM⁹⁵³

One of the key strengths of IDS RAM is that it proposes an architecture for secure data exchange and trusted data sharing and contributes to the design of enterprise architectures in commercial and industrial digitization scenarios. According to the RAM document the IDS aims to position itself as an architecture that links different cloud platforms through policies and mechanisms for secure data exchange and trusted data sharing. The central technological building block of IDS is the Connector (Figure 48) which is a dedicated software component allowing the exchange, share and process of digital content to different participants. This component also guarantees data sovereignty of the owner.

As mentioned, one of the fundamental requirements of IDS is establishing trust for data sharing and data exchange. In this context, the IDS RAM defines two basic types of trust: one based on the certification of participants and core technical components (static trust) and based on active monitoring of participants and core technical components (dynamic trust).

REFERENCE ARCHITECTURES, WITH IOT FOCUS

ISO/IEC 30141: 2018 IOT RA

The ISO/IEC CD 30141 (developed by ISO/IEC JTC 1) is a standardized IoT Reference Architecture using a common vocabulary, reusable designs, and industry best practices. It uses a top-down approach, provides a generic IoT conceptual model, a high-level system-based reference model and five architecture views: functional view, system view, user view, information view and communication view⁹⁵⁵.

⁹⁵⁵ <https://www.iso.org/standard/65695.html>

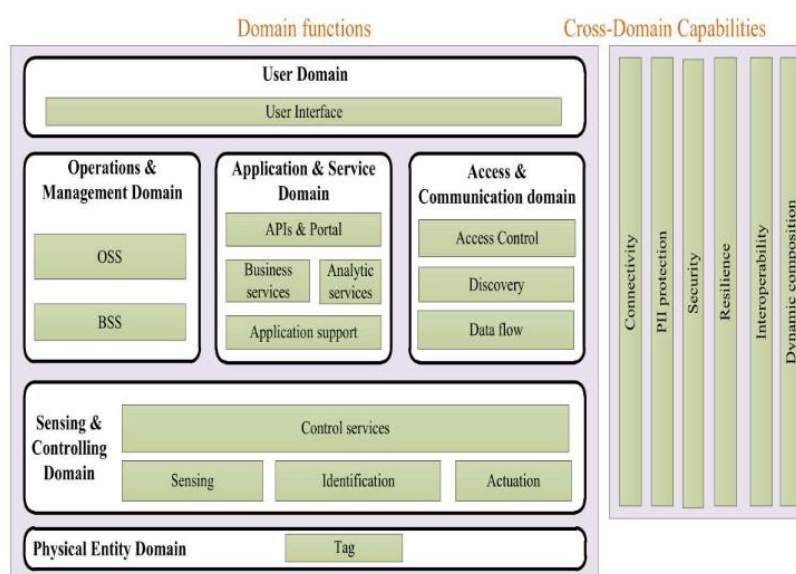


Figure 49: ISO/IEC CD 30141 (<https://www.iso.org/standard/65695.html>)

The functional view is technology-agnostic view of the functions necessary to form an IoT system and describes the dependencies among functions for support of activities described in the user view and addresses the concepts of intra and cross domain functions. The functional view of the ISO/IEC CD 30141 architecture is depicted in Figure 49.

The system view describes the generic components including devices, sub-systems, and networks required to form an IoT system. This means that the system view describes the IoT system through its physical components like subsystems, devices, and networks.

The user view focuses on how the IoT system is developed, tested, operated, and used from a user perspective and addresses the activities, roles and sub-roles and services and cross-cutting aspects.

The information view defines the information for entities on a conceptual level where data is defined as pure values without relevant or useable context.

The communication view describes the principal communications networks which are involved in IoT systems and the entities with which they connect. These are proximity, access, services, and user network.

REFERENCE ARCHITECTURES, WITH DIGITAL PUBLIC SERVICES FOCUS

EUROPEAN INTEROPERABILITY REFERENCE ARCHITECTURE (EIRA)

The [New European Interoperability Framework](#) (EIF) offers the following definition of interoperability: “The ability of organisations to interact towards mutually beneficial goals, involving the sharing of information and knowledge between these organisations, through the business processes they support, using the exchange of data between their ICT systems”.

The EIRA© is a four-view reference architecture for delivering interoperable digital public services across borders and sectors. It defines the required capabilities for promoting interoperability as a set of Architecture Building Blocks (ABBs). The EIRA© has four main characteristics:

- Common terminology to achieve coordination: It provides a common understanding of the most salient Architecture Building Blocks needed to build interoperable public services.
- Reference architecture for delivering digital public services: It offers a framework to categorise (re)usable Solution Building Blocks (SBBs) of an eGovernment solution. It allows portfolio managers to rationalise, manage and document their portfolio of solutions.
- Technology- and product-neutral and a service-oriented architecture (SOA) style: The EIRA© adopts a service-oriented architecture style and promotes ArchiMate® as a modelling notation. In fact, the EIRA© ABBs can be seen as an extension of the model concepts in ArchiMate®.
- Alignment with EIF and TOGAF: The EIRA© is aligned with the New European Interoperability Framework (EIF) (2) and complies with the context given in the European Interoperability Framework – Implementation Strategy (EIF-IS) (4). The views of the EIRA© correspond to the interoperability levels in the EIF: legal, organisational, semantic and technical interoperability which are already anchored in the National Interoperability Frameworks (NIFs) of the Member States. The EIRA© provides an additional view that lists the principles that are outlined in the new EIF. Within TOGAF® and the Enterprise Architecture Continuum, EIRA© focuses on the architecture continuum. It reuses terminology and paradigms from TOGAF® such as architecture patterns, building blocks and views. This not only assures a high level of quality but also allows architects to easily understand EIRA© and relate it to existing work.

The EIRA©⁹⁵⁶ has the objective to respond to the above needs by supporting users in the following scenarios:

- Designing: accelerate the design of eGovernment solutions that support the delivery of interoperable digital public services (across borders and sectors);
- Assessing: provide a reference model for comparing existing architectures in different policy domains and thematic areas, to identify focal points for convergence and reuse;
- Communicating and sharing: help documenting the most salient interoperability elements of complex solutions and facilitate the sharing of (re)usable solutions.
- Discovering and reusing: ease the discovery and reuse of interoperability solutions.

More specifically, the reference architecture targets the following users within public administrations of Member States or EU institutions:

- Architects, Enterprise Architects as well as Solution Architects, responsible for the design of solution architectures;
- Business analysts responsible for assessing and to study the impact of changes in the (external) environment on IT systems;
- Portfolio managers responsible for maintaining the catalogue of assets related to the design and implementation of eGovernment solutions and for making investment decisions on these assets.

⁹⁵⁶ EIRA v5.0.0 Online Documentation <https://joinup.ec.europa.eu/collection/european-interoperability-reference-architecture-eira/solution/eira/eira-v500-online-documentation>

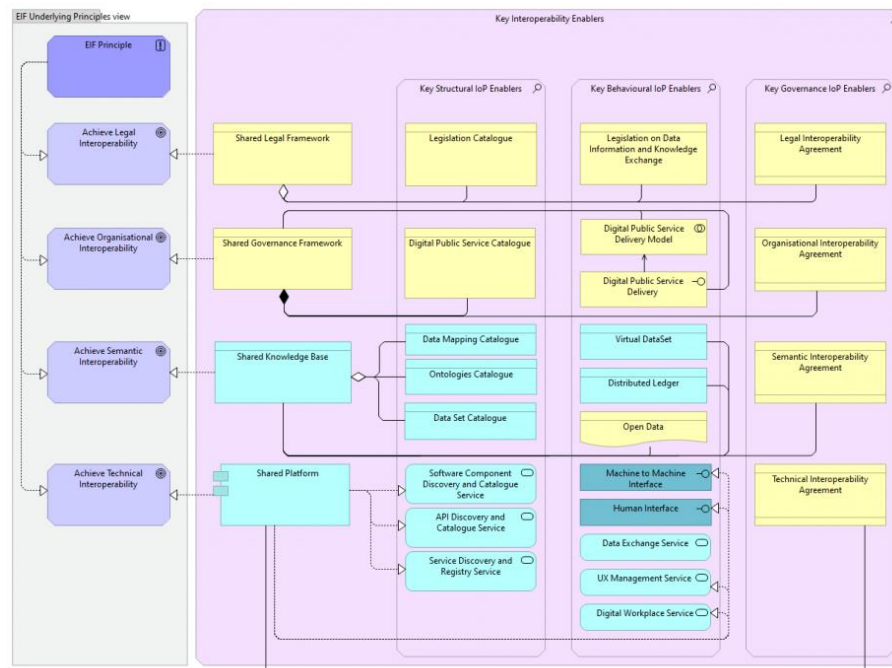


Figure 50: Key concepts of the EIRA and their relationships

Figure 50 illustrates the key concepts of the EIRA© and their relationships. The terminology is based on TOGAF®. The following list explains the different relationships depicted in Figure 50:

- The EIRA© has EIRA© Views, each EIRA© view aligns with one or more EIF Interoperability Levels
- Each EIRA© view has EIRA© Architecture Building Blocks
- The EIRA© has EIRA© Viewpoints that conform to EIRA© Views
- An EIRA© Architecture Building Block is modelled as a specialisation of a TOGAF® Architecture Building Block
- A Key Interoperability Enabler is an EIRA© Architecture Building Block, which is necessary to enable the efficient and effective delivery of public services across administrations;
- An EIRA© Architecture Building Block has interoperability requirements. An Interoperability Requirement is a statement of an interoperable need that must be realised by a system. Interoperability Requirements can be formulated for all the EIF interoperability levels: Legal Interoperability Requirements, Organisational Interoperability Requirements, Semantic Interoperability Requirements, and Technical Interoperability Requirements.
- Interoperability requirements are grouped in Interoperability Aspects. An Interoperability Aspect is an externally observable characteristic or a set of characteristics to be provided/supported by the solution that fulfils partially or internally a stakeholder interoperability need.
- An Interoperability Specification is a document containing agreed normative statements for solution building blocks used in an information exchange context. It can refer to existing standards or specifications. An Interoperability Specification realises an Interoperability Requirement.
- An EIRA© Solution Building Block is a realisation of an EIRA© Architecture Building Block and a specialization of a TOGAF® Solution Building Block
- A Solution consists of EIRA© Solution Building Blocks and TOGAF® Solution Building Blocks

The key concepts of the EIRA© are defined as follows:

- 1) EIF interoperability level: The New European Interoperability Framework (EIF) is a set of guidelines for developing public services. They cover legal, organisational, semantic and technical interoperability. Each level deserves special attention when a new European public service is established.
- 2) EIF principle: The New European Interoperability Framework outlines 12 underlying principles of European public services. These general principles of good administration are relevant to the process of establishing European public services. They describe the context in which European public services are decided and implemented. They complement one another regardless of their different natures, e.g. legal or technical. More information on the EIF interoperability levels and principles can be found in the European Interoperability Framework (EIF).
- 3) EIRA© view: The EIRA© consists of several views, including one view for each of the EIF interoperability levels. The EIRA© views contain a graphical notation of the EIRA© ontology.
- 4) EIRA© viewpoint: The EIRA© provides several viewpoints that conform to EIRA© views, the viewpoints provide a perspective with specific stakeholder's concern in mind.
- 5) Architecture Building Block: Based on the TOGAF® definition, an Architecture Building Block is an abstract component that captures architecture requirements and that directs and guides the development of Solution Building Blocks. An ABB represents a (potentially re-usable) component of legal, organisational, semantic or technical capability that can be combined with other Architecture Building Blocks. An Architecture Building Block describes generic characteristics and functionalities. Architecture Building Blocks are used to describe reference architectures, solution architecture templates or solution architectures of a specific solutions.
- 6) Solution Building Block: Based on the TOGAF® definition, a Solution Building Block is a concrete element that defines the implementation and fulfils the required business requirements of one or more Architecture Building Blocks. On the technical view, a Solution Building Block is a specific product or software component and may be either procured or developed.
- 7) Solution Architecture Template (SAT): A solution architecture template (SAT) is a specification containing including a sub-set of Architecture Building Blocks of the EIRA© and some optional Solution Building Blocks. It focuses on the most salient building blocks needed to build an interoperable solution addressing a particular business capability involving business information exchange. A solution architecture template can include additional interoperability specifications. It is usually applied within a community. Acting as a template for solutions (and their specific architectures), it guides the development of a certain kind of solutions (and their specific architectures). A solution architecture template can exist on different levels of details. For example, it can be used to describe a template for national portals offering e-services to its citizens. It can also be used to describe a template on how to securely exchange files among public administrations.
- 8) Reference Architecture: Architecture is the structure of components, their interrelationships, and the principles and guidelines governing their design and evolution over time. A reference architecture is a generalized architecture of a solution, based on best-practices, domain neutral and, occasionally, with a focus on a particular aspect. The goal of a reference architecture is reusability; it reduces the amount of work, reduces errors and accelerates the development of solutions. A reference architecture should be based in a [reference] model and in a style. The model covers the ontology of the components and their interrelationships and in the case of EIRA© it is ArchiMate®. The architecture style covers the architecture design principles and patterns and in the case of the EIRA© it is "Service Oriented Architecture" (SOA). The focus of the EIRA© is interoperability in public administrations. This definition of "reference architecture" needs to be complemented with the notion of Enterprise Architecture, which is an end-to-end generic domain neutral approach to design the architecture of an enterprise or a solution. The goal of an enterprise architecture is to align IT-related activities with the overall goal of the enterprise.
- 9) Solution Architecture: Based on TOGAF®, a solution architecture is "a description of a discrete and focused business operation or activity and how information systems / technical infrastructure supports that operation. A Solution Architecture typically applies to a single project or project release, assisting in the translation of requirements into a solution vision, high-level business and/or IT system specifications, and

a portfolio of implementation tasks”. Within the context of the EIRA©, the solution architecture describes the specific architecture of a solution. It can be derived from a solution architecture template.

- 10) Solution. A solution consists of one or more Solution Building Blocks to meet a certain stakeholder need. Within the context of the EIRA©, a solution is usually an Interoperable European Solution developed by public administrations that facilitate the delivery of electronic Public Services and cross-border exchange of information between public administrations or Citizens in support to the implementation and advancement of EU, national or local public policies.

REFERENCE ARCHITECTURES, WITH BIG DATA FOCUS

BDVA REFERENCE ARCHITECTURE

The BDV Reference Model has been developed by the European BDVA (Big Data Value Association), an industry-led organisation representing large businesses, small and medium-sized enterprises (SMEs), and research organisations in Europe. The BDVA was responsible for developing Strategic Research and Innovation Agenda (SRIA) and defining the overall goals, main technical and non-technical priorities, and a research and innovation roadmap for the European Public Private Partnership (PPP) on Big Data Value.

The BDV Reference Model addresses the main concerns and aspects to be considered for Big Data value systems. The structure of the model is separated to horizontal and vertical concerns, illustrated in Figure 51.

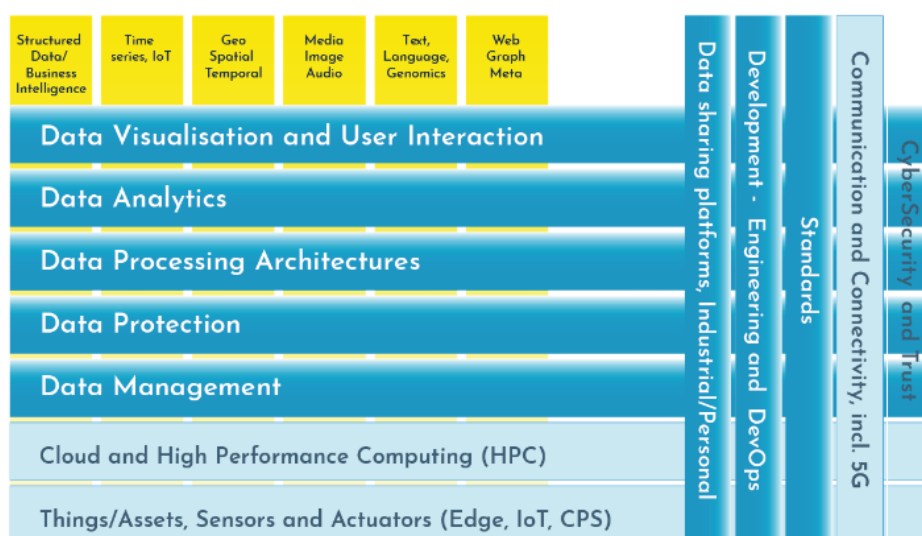


Figure 51: The BDVA reference architecture (source: www.bdva.eu)

Horizontal concerns are related to specific aspects of data processing chain, but no necessarily implying layered architecture. These are:

- Data Visualisation and User Interaction: advanced visualisation approaches that improve user experience
- Data Analytics: aspects for improving data understanding, deep learning and the meaningfulness of data.
- Data Processing Architectures: related to creating optimised and scalable architectures for analytics, with low latency delivering real-time analytics
- Data Protection: related to various privacy and anonymisation mechanisms to facilitate data protection
- Data Management: related to principles and techniques for data management

- The Cloud and High-Performance Computing (HPC): related to the fact that Big Data processing and data management implies the effective usage of Cloud and High-Performance Computing infrastructures
- IoT, CPS, Edge and Fog Computing: related to aspect of main source of Big Data which are sensor data from an IoT context and actuator interaction in Cyber-Physical Systems. Some processing is necessary to handle at the edge, to meet real-time needs

Vertical concerns address cross-cutting issues, which may affect all the horizontal concerns and may also involve non-technical aspects. These are:

- Big Data Types and Semantics: The BDV RA identified the following data types (support both the syntactical and semantic aspects of data):
 - 1) Structured data;
 - 2) Time series data;
 - 3) Geospatial data;
- Communication and Connectivity: these are necessary in providing support for Big Data
- Cybersecurity: support to maintain security and trust beyond privacy and anonymisation. This aspect has links to trust mechanisms such as blockchain technologies, smart contracts and various forms of encryption.
- Engineering and DevOps for building Big Data Value systems
- Marketplaces, Industrial Data Platforms and Personal Data Platforms (IDPs/PDPs), Ecosystems for Data Sharing and Innovation Support: Various data platforms for data sharing (like IDPs and PDPs, Research Data Platforms (RDPs), Urban/City Data Platforms (UDPs) facilitate the efficient usage of a number of the horizontal and vertical Big Data areas, most notably data management, data processing, data protection and cybersecurity.

REFERENCE ARCHITECTURES, WITH A FOCUS ON DATA SPACE

OPENDEI REFERENCE ARCHITECTURE FOR CROSS-DOMAIN DIGITAL TRANSFORMATION

OpenDei targets insights to the most relevant work in the field of Reference Architecture for building Digital Platforms to support the Digital Transformation journeys in the four sectors (manufacturing, agriculture, energy, and healthcare). The Open Dei Reference Architecture Framework (RAF) defines specifications, underlying principles, interoperability needs for the RAF.

One of the tasks of the CSA (Coordination and Support Action) OpenDei (<https://www.opendei.eu/>) is to harmonize different Digital Transformation approaches under a common Reference Architecture Framework (RAF). Main targeted sectors are Manufacturing, Energy, Healthcare and Agrifood.

The extensive use of sensors and connected devices is a common scenario in the implementation of many Digital Transformation solutions and in many industrial sectors. The huge amount of available data is able to cover many business scenarios. Data-driven pipelines and workflows management is nowadays crucial for data gathering, processing, and decision support. To cope with this complexity OPEN DEI has adopted the following 6C architecture, adapted from the one suggested by the German Industrie 4.0 initiative, and based on the following pillars (using a bottom-up reading):

- Connection, making data available from/to different networks, connecting systems and digital platforms, among several IT culture and cross organizations' boundaries, start from the capability to make data available from/to different physical and digital assets. Different devices or sensors are used to acquire a variety of IoT data, but also many systems are based on unstructured or multi-media files. Data and

information may also come from existing IT systems, using sector-specific protocols or more common standards coming from the Internet of Things (IoT) world used to realize data transfers.

- Cyber, modeling and in-memory based solutions to convert data into information, leveraging several information conversion mechanisms. Digital representations (of assets, data and information) will be then shared with upper layers of the pyramid in order to improve the self-healing properties of the overall system.
- Computing, storing and using data on the edge or on cloud. Many of modern digital platforms use a combination of cloud and edge computing models, based on driving factors for established a more centralized and powerful computation capabilities, or faster, connectivity-friendly and secure computing at the edge of the digital networked platform. The forces fueling the demand for distributing computing technologies are advancing rapidly. This will create a paradigm shift for organizations moving along new digital transformation pathways, with potential changes affecting all players in the target business ecosystem.
- Content/Context, correlating collected data for extracting information, creating a digital space for data-information continuum, not something to push out to one side of the adopted information architecture. Modern businesses need a holistic approach with the end goal of driving the data (processing) and information needs. However, exploiting data is not as straightforward. So, data needs to be acquired (captured, entered via a data pipeline) and processed with a goal and context in mind, making it information, which essentially is about processed data, before moving to the next levels.
- Community, sharing data between people and connecting stakeholders for solving collaboration needs. Networked organizations will be able to collect and share knowledge and opportunities in the widest number of sectors so that its members can make the right decisions. The community around organizations could become increasingly important to collect and share information in a push&pull fashion.
- Customization, personalizing allows to add value to data following each own user perspective and to match their expectations. Multiple strategies can make it possible to address all aspects of the end user expectations and empower an individual to progress through platform functionalities in a natural way. Democratizing access to data is a promising approach to help unlock the value of data, but even the most advanced technology is of little value if people do not embrace it. This is a lesson that many businesses have learned the hard way; in order to avoid pitfalls, it is paramount to properly understand end user expectations and build the platform from the ground up while keeping in mind that the intended audience, even within a single organization, can be very diverse and must be properly segmented and with specific and varying needs.

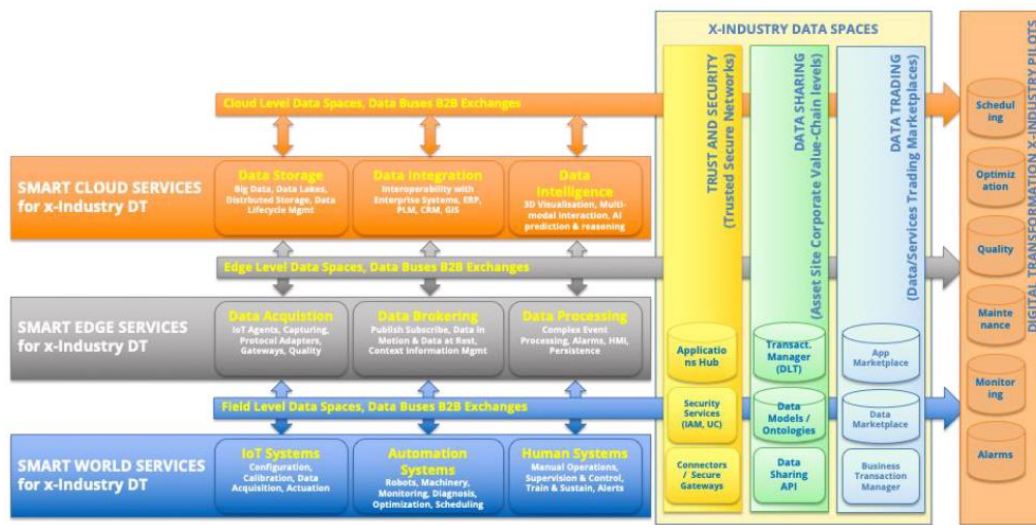


Figure 52: OpenDei reference architecture framework⁹⁵⁷

OpenDei RAF has three layers (Figure 52), bottom up namely:

- Field Level Data Spaces, it includes the Smart World Services able to collect data and support the interaction with the IoT Systems (configuration, calibration, data acquisition, actuation, etc.), Automation and Smart Assets (robots, machinery, and related operations) and Human Systems (manual operations, supervision, and control, etc.).
- Edge Level Data Spaces, it defines the typical edge operations from the data acquisition (from the logical perspective) to the data processing through the data brokering. The edge services will play a key role also for data analytics (i.e. validating and improving models for data analysis).
- Cloud Level Data Spaces, it includes data storage, data integration and data intelligence operations on the cloud. The cloud services will process big data, deploy algorithms, integrate different source platforms and services, provide advanced services such as AI prediction and reasoning.

Main orthogonal dimension named X-Industry Data Spaces, is characterized by components:

- Trust and Security,
- Data Sharing, and
- Data Trading.

Finally, all the mentioned layers serve the realization of Digital Transformation X-Industry Pilots, for enabling applications (sometimes sector specific) for supporting business scenarios.

SSI (SELF-SOVEREIGN IDENTITY) ARCHITECTURE – IMPLEMENTED IN GAIA-X

GAIA-X⁹⁵⁸ enables a federated and secure data infrastructure, whereby data are shared, with users retaining control over their data access and usage.

⁹⁵⁷ <https://www.opendei.eu/case-studies/d2-1-reference-architecture-for-cross-domain-digital-transformation/>

⁹⁵⁸ Gaia-X European Association for Data and Cloud AISBL, “Gaia-X: A Federated Secure Data Infrastructure,” 2022. [Online]. Available: <https://gaia-x.eu/>. [Accessed February 2023].

The GAIA-X solution concept⁹⁵⁹ is based on central technical requirements for the architecture of a federated, open data infrastructure: 1) Data sovereignty; 2) Use of transparently secure, open technologies; 3) Decentralized or distributed data processing; 4) Interoperability; 5) Independent and automated certification and contracting of a participant in the GAIA-X ecosystem; 6) Provision of central services that the ecosystem needs for secure and user-friendly operation and 7) Self-describing GAIA-X nodes.

To cover the most of these requirements, the GAIA-x Framework have been created to enable the transition from disjoint data & infrastructure ecosystems, to composable, interoperable & portable cross-sector data sets and services. The GAIA-X framework covers three main pillars:

1. Compliance: for a common digital governance based on European values.
2. Federation services: a mean for interoperability across federated ecosystems.
3. Data exchange: A mean to perform data exchange and anchor data contract negotiation results into the infrastructure.

For each of these pillars there are 3 types of deliverables: Functional specifications, Technical Specifications and Software. GAIA-X provides detailed documentation to understand how to belong to the GAIA-X ecosystem. A common technical documentation for these three pillars is the TRUST Framework. The Trust Framework is the set of rules that define the minimum baseline to be part of the GAIA-X Ecosystem. Those rules ensure a common governance and the basic levels of interoperability across individual ecosystems while letting the users in full control of their choices. The Trust Framework uses verifiable credentials and linked data representation to build a FAIR knowledge graph of verifiable claims from which additional trust and composability indexes can be automatically computed⁹⁶⁰.

This section of the SOTA is focused on understanding how GAIA-X allows a secure and trusted data exchange between different data sources.

GAIA-X TRUST FRAMEWORK - SSI

Self-Description (SD) is a major element of the GAIA-X architecture which describes services, participants and data assets. Participants in a GAIA-X ecosystem talk to each other in the language of Self-Descriptions⁹⁶¹. Providers use metadata to describe themselves, their service offerings, as well as the resources that their service offerings are composed of. SD uses the trustworthy Verifiable Credentials (VC), W3C standards, Decentralised Identity (DID), and associated exchange protocols to define and exchange data with its properties.

GAIA-X is developing a TRUST Framework⁹⁶⁰ and a Labelling framework⁹⁶² automating all tests and verifications needed to give a service a specific label. A label can be based on several TRUST framework criteria of which each one can include one or several Verifiable Credentials (VC).

GAIA-X provides an additional trust layer through ideally a distributed registry (Distributed Ledger Technology-DLT). DLT is used as a Verifiable Data Registry (VDR) for the secure and trustworthy share of public keys to offer cryptographic protection as an anonymous revocation registry.

⁹⁵⁹ Federal Ministry of Education and Research and Federal Ministry of Economic Affairs and Energy, “Project GAIA-X: A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem,” 2019.

⁹⁶⁰ Gaia-X European Association for Data and Cloud AISBL, “GAIA-X TRUST FRAMEWORK,” 2022.

⁹⁶¹ Gaia-X European Association for Data and Cloud AISBL, “Self-Description of Resources, Service Offerings and Participants within Gaia-X Ecosystems,” 2022.

⁹⁶² Gaia-X European Association for Data and Cloud AISBL, “GAIA-X Labelling Framework,” 2021.

To provide a trustworthy communication⁹⁶³, SSI uses SSI agent as a user interface implementing the communication protocols like DIDComm (Decentralised Identity Communication)⁹⁶⁴ or OIDC (OpenID Connect)⁹⁶⁵ for secure agent-to-agent data exchange. Therefore the agent frameworks are used to communicate and exchange DIDs and VCs with one another.

GAIA-X FEDERATION SERVICES- SSI

A Gaia-X key differentiator is the ability to establish a fully decentralised and autonomously managed ecosystem anchored in the Gaia-X Federation Services (GXFS) Identity and Trust concept, whereby each actor can interact with any other participant in a direct way. In addition to decentral Identity Management & Trust, GXFS utilises SSI in different areas like i) identity & trust; ii) Federated catalogue; iii) Compliance; iv) Data Sovereignty Services and v) Portal [6].

GAIA-X DATA EXCHANGE AND DATA PROTECTION - SSI

SSI zero-knowledge proof concepts and Attribute-Based Access Control (ABAC) enable future zero-trust and privacy protection architectures, without the disadvantages of static Role-Based Access control (RBAC) [6]. Among other things, Zero-Knowledge Proofs (ZKP) are essential cryptographic functions that allow data protection- compliant and data-saving exchange of Verifiable Credentials (VCs) in the context of SSI. With these special digital signature suites, VCs are cryptographically signed by the issuer in a special way when they are created to be able to implement the some ZKP mechanisms like

- i) Selective disclosure: Enabling to hide the information within the VC that is not required
- ii) Predicate Proofs: Help to implement the GDPR and confirm Privacy and data protection aspects simply and effectively, as information can be cryptographically verified without having to disclose the information values to a requester
- iii) Signature Blinding: allows to cryptographically hide the digital signature or the corresponding proof of an issuer by randomising the digital signature before it is passed on to an application

Private Holder Binding: allows to bind a VC cryptographically to a user and to prove this connection later without using or disclosing the user's Decentralised Identity (DID).

X-ROAD ARCHITECTURE

X-Road is open-source software and ecosystem solution that provides unified and secure data exchange between organisations. X-Road is a standardised, cohesive, collaborative, interoperable, and secure data exchange layer that gives service providers an entirely new opportunity to make themselves visible in services directed at citizens, businesses, and civil servants. Creating entities that combine many different services and data sources is easy and cost-efficient.

ECOSYSTEM

⁹⁶³ Gaia-X European Association for Data and Cloud AISBL, "Gaia-X secure and trustworthy ecosystems with Self Sovereign Identity," Brussels, 2022.

⁹⁶⁴ DIDComm, "Share protocols built on DIDComm," 2023. [Online]. Available: <https://didcomm.org/>. [Accessed February 2023].

⁹⁶⁵ OpenID, "Welcome to OpenID Connect," 2023. [Online]. Available: <https://openid.net/connect/>. [Accessed February 2023].

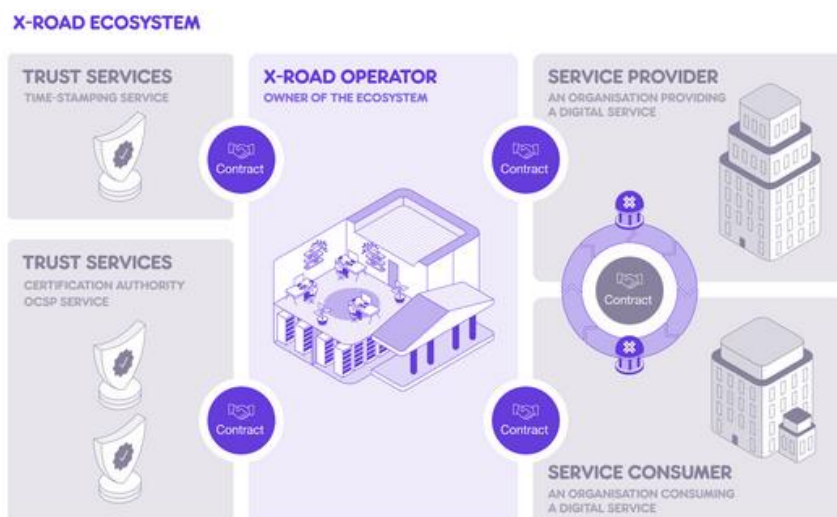


Figure 53: X-Road ecosystem

An X-Road ecosystem (Figure 53) is a community of organizations using the same instance of the X-Road software for producing and consuming services. The owner of the ecosystem, the X-Road Operator, controls who are allowed to join the community, and the owner defines regulations and practices that the ecosystem must follow. The ecosystem may be nationwide, or it may be limited to organizations meeting specific criteria, e.g., clients of a commercial service provider. Technically, the X-Road software does not set any limitations to the size of the ecosystem or the member organizations.

As the owner of the X-Road ecosystem, the Operator is responsible for all aspects of the operations. The responsibilities include defining regulations and practices, accepting new members, providing support for Members, and operating the central components of the X-Road software. X-Road Members are organizations that have joined the ecosystem and produce and/or consume services with other Members. A Member organization can be a service provider, a service consumer, or both. Organizations can become Members of an ecosystem by completing the onboarding process defined by the Operator. Also, members need to have access to the technical component that is required for exchanging messages via X-Road, the Security Server. A functioning X-Road ecosystem requires two types of trust services: 1) time-stamping authority (TSA) and 2) certification authority (CA). Trust Service Providers are organizations providing these services. Trust Service Providers may be commercial third parties, or the services can be provided and maintained by the X-Road Operator too.

TRUSTED NETWORK

Even if X-Road software is open-source, joining an X-Road ecosystem requires going through an onboarding process. During the process, the identity of each organization and technical access point is verified using certificates that are issued by a trusted Certification Authority (CA). The identities are maintained centrally, but all the data is exchanged directly between a service consumer and a service provider.

Message routing is based on organization and service level identifiers that are mapped to physical network locations of the services by X-Road. All the evidence regarding the data exchange is stored locally by the data exchange parties, and no third parties have access to the data. Time-stamping and digital signature together guarantee non-repudiation of the data sent via X-Road. The logs provided by X-Road can be used in a court proceeding as evidence.

AUTHORIZATION FRAMEWORK

X-Road implements an authorization framework that is used to manage access rights to services. Access rights management is based on the organization and service level identifiers.

The key idea of X-Road is that each service provider owns its data and is responsible for managing access rights of its services. In other words, publishing service to X-Road does not mean that the service is automatically accessible to all X-Road member organizations. Usually, access rights are granted on the information system level – a service provider grants a specific information system access to a service.

MONITORING AND REPORTING

X-Road provides monitoring and reporting capabilities that can be used to collect operational reporting data and technical monitoring information from the ecosystem. The information can be used to measure the usage of individual services, understand dependencies and relationships between different information systems and services, monitor service health, monitor used X-Road software versions, etc. Each X-Road member organization can access its own data, whereas the X-Road operator can access all the members' data.

CROSS-BORDER DATA EXCHANGE

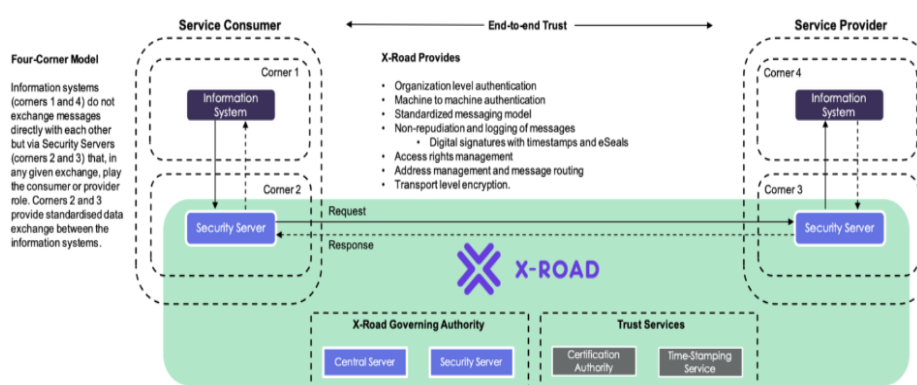


Figure 54: Cross-border data exchange [NIIS – X-Road and eDelivery]

X-Road provides built-in support for cross-border data exchange through federation, which means joining together two X-Road ecosystems. Members of the federated ecosystems can publish and consume services with each other as if they were members of the same ecosystem.

It is possible to create federation connections with multiple ecosystems, but transitive federation relationships are not supported. An ecosystem does not have a federation relationship with another ecosystem that it's not directly federated with.

EU AI FRAMEWORK

To boost value-driven adoption of AI across European industrial sectors, the public sector, and society, strategy and mechanism that aligns user value and industrial offerings with research excellence in AI are needed. The creation of new AI Public Private Partnership (PPP) aims at the coordination of all stakeholders in the emerging European AI Ecosystem, encompassing end-users, small, medium and large industrial players, start-ups, research institutes and academia, education and training, policymakers and regulators, entrepreneurs and investors. It will create synergies between different communities to optimise the impact of European investments in AI.

The European Commission's Coordinated Plan on Artificial Intelligence calls for the development of an industrially led AI PPP triggered by the Big Data Value Association (BDVA) and the European Robotics Association

(euRobotics) through the joint action of their respective cPPPs. The AI PPP recognises that the full value of AI comes when it unifies information and motion, digital and physical, data and robotics.

The European AI Framework is based on:

- **European Fundamental Rights, Principles, and Values:** All technologies of the AI Value Chain will follow core European rights, principles, and values¹¹, such as respect for human dignity and human rights, freedom, democracy, equality and the rule of law. This European core will serve as high-level requirements for all AI innovations and build trust in AI transformations.
- **Value-Driven for Business, Society, and People:** To foster adoption by a wide range of industries, the public sector, society and citizen, value-driven AI will focus on markets and applications. The value realisation will have a specific focus on the industrial challenges of sectors and domains that are strong in and of importance for Europe.
- **Policy, Regulation, Certification, and Standards:** The AI PPP will contribute to the policy and regulatory debate about the non-technical aspects of AI-driven value creation. By promoting certification and standards in AI, the adoption and scaling of AI technologies will be supported.

AI Ecosystem Enablers identified within the framework include:

- **Business and Technical Skills and Knowledge Exchange:** As traditional industry sectors undergo an AI transformation, so too must their workforces. The specialisation required by AI engineers will deepen as the sophistication of leading-edge tools and algorithms increases. The skills for general workers will become broader with increased AI fluency needed including numeracy and statistics. The ability to judge bias in both data and algorithms will necessitate transdisciplinary training for knowledge workers. The delivery of AI skills to SMEs will also be important.
- **Systems Development and Data for AI:** Interdisciplinary approaches which provide strategies, processes and tools for designing and managing complex AI solutions and systems. Next generation AI technology seeks to realise secure and trusted data sharing between diverse participants so they can exchange and gain access to new and broader data within a clear legal framework. This will be facilitated by data platforms for sharing open data, industrial data and personal data, giving full control to the data providers and data subjects over their data, encoring compliance with legal and ethical requirements, and ensuring that full provenance information is maintained.
- **Experimentation and Innovation:** The partnership will stimulate experimentation environments where companies, and more specifically SMEs, can have their AI related services and products quickly piloted and tested, before going to market, for example by using AI Digital Innovation Hubs, where they can access the necessary infrastructure, technical support, skills, data, etc. including incubation and acceleration services, to enable companies to rapidly develop new businesses based on AI technologies, applications and models.

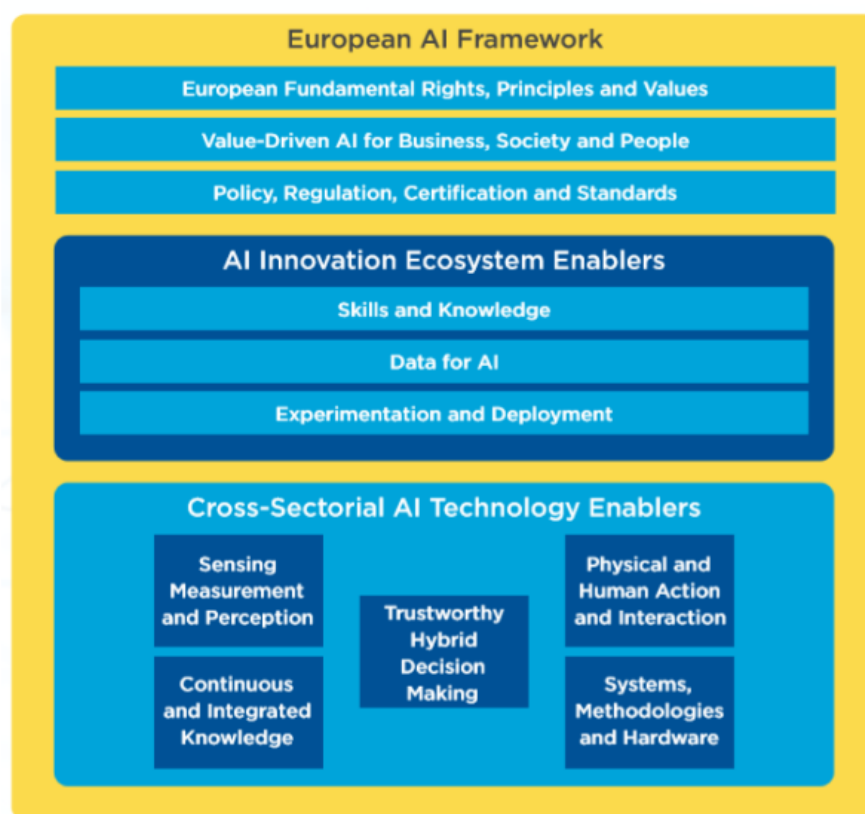


Figure 55: EU AI Framework

Cross-Sectorial AI Technology Enablers:

- **Acquisition & Sensing:** Data acquisition and sensing technologies will provide methods for accessing, converting and aggregating raw data and signals measuring real-world physical parameters into data assets that can be used as input to subsequent analysis steps.
- **Data Processing and Analysis:** Data processing and analysis will make the raw data acquired amenable to use in decision-making. This involves transforming, cleaning, storing, sharing, modelling, synthesising and extracting insights.
- **Decision-Making:** AI-driven applications and activities need access to data, its analysis, and the tools needed to integrate the data analysis within the system or activity. AI usage in business or autonomous decision-making can enhance competitiveness through the reduction of costs, increased added value, or other parameters that can be measured against existing performance criteria.
- **Physical and Human Action and Interaction:** The value of AI solutions is greatest where AI-driven insights can be used to assist humans in making complex decisions or where it directs machines to navigate and physically manipulate objects in complex and changing work environments.

TRUSTEE LANDSCAPE

TRUSTEE Landscape consists of projects and initiatives related to the objectives and issues defined by the TRUSTEE project. This section provides an overview of the identified activities still active at the time of the TRUSTEE project start.

RELATED EU PROJECTS

One of the relevant goals of the state-of-the-art analysis conducted in the initial phase of the project was to identify EU projects related to TRUSTEE. Our main focus was on the ongoing projects since we wanted to gain insight into the current landscape of EU projects that could potentially form a collaboration with the TRUSTEE project. However, later in this chapter we also included some of the recently completed EU projects which we considered especially interesting due to the achieved results related to the scope of TRUSTEE.

We gathered information on relevant related projects by iteratively asking all project partners to provide inputs on relevant projects that they know of. Inputs were collected in an excel file stored in the project repository. After the list of the proposed related project was formed, we analysed each of the projects and evaluated several aspects, such as the maturity of the project results, participating project partners (potentially indicating facilitation of collaboration), and alignment and overlapping of the topics and goals with the TRUSTEE project. Finally, 22 of them were in this document, as they demonstrated relevance for potential future collaborations and alignment with TRUSTEE project goals. For this reason, special focus was on the ongoing related projects identified at the time of conducting analysis (September 2022 – February 2023) that were selected as a relevant part of the TRUSTEE landscape.

TRUSTS

Full project name: Trusted Secure Data Sharing Space

Website: <https://www.trusts-data.eu/>

Start/End date: 01 January 2020 to 31 December 2022

Programme: H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT)

Summary of project objectives:

The development of a data market and data economy is hampered by the lack of trusted and secure platforms and privacy-aware analytics methods for the secure sharing of personal data and proprietary/commercial/industrial data, primarily limiting data sharing to open data. If various technical standards, quality levels, and legal considerations are permitted to uncontrollably diverge, this trend will continue.

TRUSTS' emphasis on creating a platform based on the knowledge gained from two significant national projects while enabling the adoption of future platforms will ensure trust in the idea of data markets as a whole. When examining the moral and legal considerations that apply to the entire data value chain, from data providers to consumers, the TRUSTS platform will act independently and as a platform federator, i.e. it will

(a) Set up a fully operational and GDPR-compliant European Data Marketplace for personal related data and non-personal related data targeting both personal and industrial use by leveraging existing data marketplaces (International Data Space and Data Market Austria) and enriching them with new functionalities and services.

(b) Demonstrate and realise the potential of the TRUSTS Platform in 3 use cases targeting the industry sectors of corporate business data, specifically in the financial and telecom operator industries while ensuring it is supported by a viable, compliant and impactful governance, legal and business model.

To create a European Data Market based on secure and trustworthy data exchanges, the TRUSTS consortium brings together technology providers that are already deeply involved in major national data market projects. This integration will be tested in practice by 6 companies, including 2 data providers, addressing 3 different use-cases. The TRUSTS technology and use-cases will be accompanied by business and legal and ethical considerations, which will ensure that the results of the project are sustainable beyond its duration.

TRAPEZE

Full project name: TRAnsparency, Privacy and security for European citiZEns

Website: <https://trapeze-project.eu/>

Start/End date: 01 September 2020 to 31 August 2023

Programme: H2020-EU.3.7. - Secure societies - Protecting freedom and security of Europe and its citizens

Summary of project objectives:

In order to enable citizens to actively contribute to the cyber resilience of the common European data space, the EU-funded TRAPEZE project will develop technologies. It will provide the public with the necessary tools and information to manage their security and privacy. It will also create technologies to enforce the integrity and non-repudiation of citizens' data usage policies and processing across data sources. Overall, the project will pave the way for the practical application of cutting-edge technologies like blockchain and linked data.

AI-SPRINT

Full project name: Artificial Intelligence in Secure PRiVacy-preserving computing coNTinuum

Website: <https://www.ai-sprint-project.eu/>

Start/End date: 01 January 2021 to 31 December 2023

Programme: H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT)

Summary of project objectives:

AI-SPRINT, a project funded by the EU, creates robust tools for AI applications in computing systems that enable a finely tuned tradeoff between performance and AI model accuracy, incorporating security and privacy from the beginning of the design process. The project offers mechanisms for automatic deployment and dynamic reconfiguration, as well as programming models and specialized building blocks.

TEHDAS

Full project name: Towards European Health Data Space

Website: <https://tehdas.eu>

Start/End date: 01 February 2021 to 31 August 2023

Programme: Co-funded by the Health Programme of the European Union

Summary of project objectives:

The joint action Towards the European Health Data Space (TEHDAS) aids EU member states and the European Commission in creating and promoting ideas for the secondary use of health data to advance public health, medical research, and technological advancement in Europe.

Better healthcare services and individualized care for patients are among the advantages, as are the advancement of innovations like the creation of new medicines and the promotion of knowledge-based policy-making.

TRUST-AI

Full project name: Transparent, Reliable and Unbiased Smart Tool for AI

Website: <http://www.trustai.eu/>

Start/End date: 1 October 2020 to 31 March 2025

Programme: H2020-EU.1.2. - EXCELLENT SCIENCE - Future and Emerging Technologies (FET)

Summary of project objectives:

The EU-funded TRUST-AI project is involving human intelligence in the discovery process. It will employ 'explainable-by-design' symbolic models and learning algorithms and adopt a human-centric, 'guided empirical' learning process that integrates cognition. The project will design TRUST, a trustworthy and collaborative AI platform, ensure its adequacy to tackle predictive and prescriptive problems and create an innovation ecosystem in which academics and companies can work independently or together.

ENCRYPT

Full project name: A Scalable And Practical Privacy-Preserving Framework

Website: To be defined, information about project available at: <https://cordis.europa.eu/project/id/101070670>

Start/End date: 1 July 2022 to 30 June 2025

Programme: HORIZON-AG - HORIZON Action Grant Budget-Based

Summary of project objectives:

ENCRYPT will develop a scalable, practical, adaptable privacy-preserving framework, allowing researchers and developers to process data stored in federated cross-border data spaces in a GDPR-compliant way. Within this framework, a recommendation engine for citizens and end-users will be developed, providing them with personalised suggestions on privacy preserving technologies depending on the sensitivity of data and the accepted trade-off between the degree of security and the overall system performance. The ENCRYPT framework will be designed taking into consideration the needs and preferences of relevant actors, and will be validated in a comprehensive, 3-phase validation campaign, comprising i) in-lab validation tests, ii) use cases provided by consortium partners in three sectors, namely the health sector, the cybersecurity sector, and the finance sector, that include cross-border processing of data, and iii) external use cases including privacy preserving computations on federated medical datasets.

OMEGA-X

Full project name: Orchestrating an interoperable sovereign federated Multi-vector Energy data space built on open standards and ready for GAia-X

Website: <https://omega-x.eu/>

Start/End date: 1 May 2022 to 30 April 2025

Programme: HORIZON.2.5 - Climate, Energy and Mobility

Summary of project objectives:

"The aim of OMEGA-X is to implement a data space (based on European common standards), including federated infrastructure, data marketplace and service marketplace, involving data sharing between different stakeholders and demonstrating its value for real and concrete Energy use cases and needs, while guaranteeing scalability and interoperability with other data space initiatives, not just for energy but also cross-sector.

The proposed concept and architecture heavily rely on the approaches adopted by IDSA, GAIA-X, FIWARE, BDVA/DAIRO and SGAM as major EU references regarding data spaces. It will pursue the GAIA-X label, which ensures highest standards on protection, security, transparency, openness and trust, avoids vendor lock-in and restricted to EU countries.

- Federated infrastructure for data ingestion. There are a lot of independent platforms for data ingestion/storage, open and private. The goal is to define the minimum interoperability and federation requirements needed for these platforms to adhere to the Energy Data Space and be able to share data in a trusted and secure way.

- Data Space Marketplaces. This is the common ground where data, which is already harmonized semantically, is indexed, and referenced, maintaining always the required standards of identity, trust and sovereignty. Using the data space as baseline, a marketplace is implemented for stakeholders to share, use and monetize data and services. Data/service providers will be able to advertise their data/services, and data/service users will be able to discover multiple data sets and services."

I4TRUST

Full project name: Incubator of Trusted B2B Data Sharing ecosystems of collaborating SMEs linked to Digital Innovation Hubs

Website: <https://i4trust.org/>

Start/End date: 1 October 2020 to 30 September 2023

Programme: H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT)

Summary of project objectives:

"i4Trust is building a sustainable ecosystem where companies will be able to create innovative services by means of breaking "data silos" through sharing, re-using and trading of data assets.

The EU-funded i4Trust project will contribute to the breaking of 'data silos' that isolate interoperability between information systems and stimulate the sharing, reusing and trading of data assets. It will launch an incubator of trusted B2B data sharing ecosystems of collaborating SMEs connected to digital innovation hubs (DIHs), allowing them to share data and encouraging them to adopt proven successful tools such as iSHARE, CEF building blocks and FIWARE data models. "

AITHENA

Full project name: AI-based CCAM: Trustworthy, Explainable, and Accountable

Website: To be defined, project information available at: <https://cordis.europa.eu/project/id/101076754>

Start/End date: 1 November 2022 to 31 October 2025

Programme: HORIZON.2.5 - Climate, Energy and Mobility

Summary of project objectives:

"Connected and Cooperative Automotive Mobility (CCAM) solutions have emerged thanks to novel Artificial Intelligence (AI) which can be trained with huge amounts of data to produce driving functions with better-than-human performance under certain conditions.

AITHENA will contribute to build Explainable AI (XAI) in CCAM development and testing frameworks, researching three main AI pillars: data (real/synthetic data management), models (data fusion, hybrid AI approaches), and testing (physical/virtual XiL set-ups with scalable MLOps)."

LAGO

Full project name: Lessen Data Access And Governance Obstacles

Website: To be defined, project information available at: <https://cordis.europa.eu/project/id/101073951>

Start/End date: 1 November 2022 to 31 October 2024

Programme: HORIZON.2.3 - Civil Security for Society

Summary of project objectives:

LAGO will deliver the foundation for a trusted EU FCT Research Data Ecosystem (RDE) to address the so-called "Data Issue" in the FCT research landscape, i.e. the lack of domain-specific data in sufficient quality and quantity to enable appropriate training and testing of the developed methods, platforms and tools. LAGO will be instrumental in identifying common barriers and subsequently providing the structural, governance and technical foundations to foster and innovate data-oriented research collaboration among LEAs, security practitioners, relevant EU agencies, academic and industry researchers, policy makers and regulators. For this purpose, LAGO will develop an evidence-based and validated multi-actor Reference Architecture for the FCT RDE for these actors to deposit, share and co-create data and tools for FCT research purposes based on common rules, protocols, standards and instruments in a trusted and secured environment. The envisaged Reference Architecture and accompanying governance framework will be based on the design principles of decentralisation, data sovereignty, data quality, openness, transparency and trust and comply with EU values and principles on data protection, privacy and ethics. The Reference Architecture will be accompanied by a TRL-7 Reference Implementation of added-value technological tools to ensure practical realisation of the Reference Architecture as multiple data spaces and across the full range of concrete usage scenarios. A Roadmap will finally provide the consolidated rules, conditions and considerations for the actual deployment of the EU FCT RDE. The ultimate ambition of LAGO is to go beyond the creation of a common repository in order to innovate the FCT data-oriented research sphere by creation the crucial foundations for the sustainable, safe and trusted creation, co-creation, sharing and maintenance of training and testing datasets for the FCT research domain.

ENERSHARE

Full project name: European common EneRgy dataSpace framework enabling data sHaring-driven Across- and beyond- eneRgy sERVICES

Website: <https://enershare.eu/>

Start/End date: 1 July 2022 to 30 June 2025

Programme: HORIZON.2.5 - Climate, Energy and Mobility

Summary of project objectives:

"ENERSHARE will deliver a Reference Architecture for a European Energy Data Space, which hybridizes SGAM with IDSA and GAIA-X architectures, by bringing data value chain perspective into the energy one, and evolve interoperability, trust, data value and governance building blocks to TRL 6-7 IDSA-compliant ones, adapt them to energy sector, and deploy:

- across-energy and cross-sector data enhancement technology enablers and standardizable interfaces and open APIs by leveraging on open Standards (e.g. ETSI Context Broker) and ontologies (e.g. SAREF
- trust-related connectors, to ensure privacy, confidentiality, cybersecurity-preserving trust, sovereignty and full control of data
- Blockchain/Smart contract-enriched marketplace for data versus energy assets/services coordination, sharing, exchange, and beyond financial compensation
- cross-value chain value-added services and Digital Twins, by leveraging on privacy-preserving federated learning
- integrate and deploy them within a Reference Implementation of a European Energy Data Space, which will be demonstrated along 7 pilots and 11 intra-electricity, intra-energy and beyond energy use cases
- co-design SSH-based consumer-centric business models for energy data sharing enabling data beyond-financial value creation and spreading along value chain
- prepare the ground for the European Energy Data Space setup, through alignment with EU-level relevant initiatives (GAIA-X, IDSA, BDVA, ETIP SNET, BRIDGE), contributing to Data Space standardization and boosting a level playing field for data sharing."

STELAR

Full project name: Spatio-TEmporal Linked data tools for the AgRi-food data space

Website: To be defined, project information available at: <https://cordis.europa.eu/project/id/101070122>

Start/End date: 1 September 2022 to 31 August 2025

Programme: HORIZON.2.4 - Digital, Industry and Space

Summary of project objectives:

STELAR will design, develop, evaluate, and showcase an innovative Knowledge Lake Management System (KLMS) to support and facilitate a holistic approach for FAIR (Findable, Accessible, Interoperable, Reusable) and AI-ready (high-quality, reliably labeled) data. The STELAR KLMS will allow to (semi-)automatically turn a raw data lake into a knowledge lake. This is achieved by (1) enhancing the data lake with a knowledge layer, and (2) developing and integrating a set of data management tools and workflows. The knowledge layer will comprise: (a) a data catalog offering automatically enhanced metadata for the raw data assets in the lake, and (b) a knowledge graph that semantically describes and interlinks these data assets using suitable domain ontologies and vocabularies. The provided tools and workflows will offer novel functionalities for: (a) data discovery and quality management; (b) data linking and alignment; and (c) data annotation and synthetic data generation.

DATA CELLAR

Full project name: Data hub for the Creation of Energy communities at Local Level and to Advance Research on them

Website: <https://datacellarproject.eu/>

Start/End date: 1 June 2022 to 30 November 2025

Programme: HORIZON.2.5 - Climate, Energy and Mobility

Summary of project objectives:

DATA CELLAR aims therefore to create a public energy dataspace that will support the creation, development and management of LECs in EU. Such dataspace will be easy to be populated (also via an innovative rewarded private metering approach) and easy to interact with, also guaranteeing a smooth integration with other EU energy dataspace and providing to LEC stakeholders services and tools for developing their activities. DATA CELLAR will implement a collaborative platform providing an interoperable, modular, and secure energy data space capable of delivering access to datasets, Decision Support Tool and Artificial Intelligence (AI) models to serve and support the spread LECs. In the framework of the project DATA CELLAR and its services will be used by representatives of nine validation cases around EU and with different level of maturity promoted by different type of actors such as public authorities, DSOs, aggregators etc.. DATA CELLAR will also used to drive transversal activities such as EU energy scenarios and market planning (EDF, NODES). DATA CELLAR will facilitate data collection/sharing from both LEC stakeholders, researchers and tenants also targeting the development of a marketplace and a tokenization rewarding approach that can stimulate such actors to both use and populate the data space. At this purpose and also considering data interaction with the other EU existing energy dataspace/initiatives (GAIA-X, OPEN DEI...) towards the creation of a federated EU Energy Data Space, DATA CELLAR will also study ethical, regulatory, cybersecurity and governance aspects for data handling/sharing.

USAGE

Full project name: Urban Data Spaces for Green dEal

Website: To be defined, project information available at: <https://cordis.europa.eu/project/id/101059950>

Start/End date: 1 August 2022 to 31 July 2025

Programme: HORIZON.2.6 - Food, Bioeconomy Natural Resources, Agriculture and Environment

Summary of project objectives:

"USAGE (Urban Data Space for Green Deal) aims to provide solutions and mechanisms for making city-level environmental and climate data available to everyone based on FAIR principles.

USAGE will provide innovative governance mechanisms, consolidated arrangements, AI-based tools and data analytics to share, access and use city-level data from Earth Observation (EO), Internet of Things (IoT), authoritative and crowd sources, leveraging on standards for data and service interoperability.

USAGE wants to become a decentralized infrastructure for trustworthy data collection, processing and exchange based on commonly agreed principles, facilitating the combination of heterogeneous data for policy analysis.

USAGE will validate its solutions in four diverse pilot areas located in four different countries, focusing also on the reusability of the solutions in other urban areas."

TERMINET

Full project name: nexT gEneRation sMart INterconnectEd IoT

Website: <https://terminet-h2020.eu/>

Start/End date: 1 November 2020 to 31 October 2023

Programme: H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT)

Summary of project objectives:

The vision of TERMINET is to provide a novel next generation reference architecture based on cutting-edge technologies such as SDN, multiple-access edge computing, and virtualisation for next generation IoT, while introducing new, intelligent IoT devices for low-latency, market-oriented use cases. TERMINET's primary intention is to bring (more efficient and accurate) decisions to the point of interest to better serve the final user targeting at applying distributed AI at the edge by using accelerated hardware and sophisticated software to support local AI model training using federated learning. Our solution aspires to reduce the complexity of the connecting vast number of heterogeneous devices through a flexible SDN-enabled middleware layer. It also aims to design, develop, and integrate novel, intelligent IoT devices such as smart glasses, haptic devices, energy harvesting modules, smart animal monitoring collars, AR/VR environments, and autonomous drones, to support new market-oriented use cases. Great expectation of the proposal is to foster AR/VR contextual computing by demonstrating applicable results in realistic use cases by using cutting-edge IoT-enabled AR/VR applications. By designing and implementing an IoT-driven decentralised and distributed blockchain framework within manufacturing, TERMINET aims to support maintenance and supply chain optimisation. Our solution intends to apply a vertical security by design methodology by meeting the privacy-preserving and trust requirements of the NG-IoT architecture. To foster standardisation activities for the IoT ecosystem, TERMINET will provide novel disruptive business models. For the evaluation of its wide applicability, TERMINET will validate and demonstrate six proof-of-concept, realistic use cases in compelling IoT domains such as the energy, smart buildings, smart farming, healthcare, and manufacturing.

OASEES

Full project name: Open Autonomous programmable cloud appS& smart EdgE Sensors

Website: <https://oasees-project.eu/>

Start/End date: 1 January 2023 to 31 December 2025

Programme: HORIZON.2.4 - Digital, Industry and Space

Summary of project objectives:

" The massive increase in device connectivity and generated data has resulted in the proliferation of intelligent processing services to create insights and exploit data in a multi-modal manner. Currently, the most powerful data processing operates in a centralized manner at the cloud, which provides the ability to scale and allocate resources on demand and efficiently. Centralized processing and cloud hosting, bound and limit their services and applications to operate in a resource restricted manner, relying usually on large single entities to provide, i) Authentication, ii) Data storage, iii) Data processing, iv) Connectivity, v) Vendor-locked environments for development and orchestration. This significantly limits the user from its data governance and even identity management. Similarly, existing solutions for edge device authentication require a centralized entity to trust them and authenticate them, rendering a non-portable identification paradigm. OASEES aims to create an open, decentralized, intelligent, programmable edge framework for Swarm architectures and applications, leveraging the Decentralized Autonomous Organization (DAO) paradigm and integrating Human-in-the-Loop (HITL) processes for efficient decision making. The OASEES vision is to provide open tools and secure environments for swarm programming and orchestration for numerous fields, in a completely decentralized manner. An important aspect in this process is identification and identity management, in which OASEES targets the implementation of a portable and privacy

preserving ID federation system, for edge devices and services, with full compliance and compatibility to GAIA-X federation and IDSA trust directives and specifications. This situation solidifies the need for an integrated enabler framework tailored to the edge's extreme data processing demands, using different edge accelerators, i.e. GPU, NPU, SNN and Quantum."

TEADAL

Full project name: Trustworthy, Energy-Aware federated DAta Lakes along the computing continuum

Website: <https://www.teadal.eu/>

Start/End date: 1 September 2022 to 31 August 2025

Programme: HORIZON.2.4 - Digital, Industry and Space

Summary of project objectives:

"Data analytics is one of the main cornerstones in many enterprise architectures and the data lake paradigm is more and more adopted to assist organizations in taking reliable, accurate, and fast decisions. Although the initial approaches to address these issues saw the data lakes as the evolution of data warehouses to be implemented on-premises, cloud providers are nowadays including in their offerings platforms able to setup and run them. Nevertheless, the increasing amount of data generated at the edge and the need to enable the data sharing among organizations are posing new challenges in terms of performances, energy efficiency, and privacy/confidentiality which can be properly addressed with data lakes which are deployed along the whole computing continuum as well as building a federation of such data lakes.

The ambition of TEADAL is to provide key cornerstone technologies to create stretched data lakes spanning the cloud-edge continuum and multi-cloud, providing privacy, confidentiality, and energy-efficient data management. The TEADAL data lake technologies will enable trusted, verifiable and energy efficient data flows, both in a stretched data lake and across a trustworthy mediatorless federation of them, based on a shared approach for defining, enforcing, and tracking privacy/confidentiality requirements balanced with the need for energy reduction."

KRAKEN

Full project name: Brokerage and market platform for personal data

Website: <https://www.krakenh2020.eu/>

Start/End date: 1 December 2019 to 30 November 2022

Programme: H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT)

Summary of project objectives:

"KRAKEN Secure and privacy preserving platform (broKeRage And marKEt platform for persoNal data) aims to bring personal data sharing and trading at a level of maturity that does not yet exist, by leveraging on:

i) the emerging paradigm of self-sovereign identity built upon a stack of distributed ledger technologies (multi-ledger) which ensures future compatibility with different specific blockchain implementations for identity management. It will provide a decentralized user-centric approach on personal data sharing and proving that it can incorporate the trust and security assurance levels deriving claims from national identity schemas (eIDAS-compliant);

- ii) tested data marketplace technologies which support data sharing as well as aggregated data sharing;
- iii) A set of different data protection techniques based on advanced crypto tools (P/F/HE, FE, MPC, ABE...) coupled with privacy preserving (AI/ML) analytics, featuring management of privacy / utility trade-offs and metadata privacy.
- iv) The project will provide market-ready tools and services with industrial strength and suitable privacy metrics that will be conveyed to data subjects with high usability.

Taking advantages of the emergent models and lessons learned from previous experiences, including existing and mature cloud-based personal data platforms CREDENTIAL and MHMD, as well as Streamr marketplace services, Kraken proposes an unprecedented approach, which creates an alternative to mainstream paradigms while fully granting the privacy and self-sovereignty of the data subject. KRAKEN enables advanced, convenient data sharing control relying on innovative end-to-end encryption and use of sophisticated proxy cryptography schemes grants data subjects with unprecedented control over their data. Furthermore, this ensures that even the cloud provider (data processor) cannot access the data in plain-text and hence protect access to personal data. KRAKEN involves 10 partners and has a duration of 3 years."

AI4EUROPE

Full project name: Europe's AI-on-Demand Platform

Website: <https://www.ai4europe.eu/>

Start/End date: 1 July 2022 to 31 December 2025

Programme: HORIZON.2.4 - Digital, Industry and Space

Summary of project objectives:

AI4Europe builds on the work of AI4EU and multiple supporting projects (ICT-48/ICT-49), creating an open, impartial, and collaborative Platform, built by the European research community according to their needs. Equipped with the necessary hardware, the Platform will offer interoperable services, data, and tools from several related communities and provide solutions to facilitate research productivity, reproducibility, and collaboration. AI4Europe will establish and support mechanisms to foster exchange between academia and industry and ensure the Platform reaches out to and engages with the next generation of researchers and those in widening countries. The project will develop and implement a business model that will ensure the long-term technical and financial structures providing sustainability for the Platform beyond the lifetime of the project. AI4Europe will support the community to create a tool that will help position Europe as the place where the very best AI research is conducted.

ENRICHMYDATA

Full project name: Enabling Data Enrichment Pipelines for AI-driven Business Products and Services

Website: <https://enrichmydata.eu/>

Start/End date: 1 October 2022 - 30 September 2025

Programme: HORIZON.2.4 - Digital, Industry and Space

Summary of project objectives:

enRichMyData provides a novel paradigm for building rich, high-quality and valuable datasets to feed Big Data Analytics and AI applications. The paradigm consists in facilitating the specification and scalable execution of data enrichment pipelines, with a focus on supporting various data enrichment operations such as discovery, understanding, selection, cleaning, transformation, integration of Big Data from a variety of sources. enRichMyData makes this paradigm easily accessible to a wide set of large and small organizations that encounter difficulties in the process of delivering suitable data to feed their data analytics solutions due to the lack of suitable tools and expertise to support cost- effective and energy-efficient management of data enrichment pipelines. enRichMyData delivers a software toolbox consisting of tools and infrastructure services for setting-up, deploying, executing and managing data enrichment pipelines. The main exploitation targets the operation and monetization of the toolbox primarily in European markets. enRichMyData validates its plan through a strong selection of complementary business cases offered by commercially- focused organizations targeting development of novel products in a wide range of domains, including digital marketing, Industry 4.0 medical devices maintenance, public procurement, innovation ecosystems, and mineral processing. The balanced consortium consists of 13 partners from 11 countries. With a mixture of R&D/technology providers and business case providers, enRichMyData gathers five large companies, three SMEs, two research institutes, and three universities.

INITIATIVES

iSHARE

Web site: <https://ishare.eu/>

Short description: iSHARE is the European standard for and trust network of international business data sharing in a sovereign way, governed by iSHARE Foundation. Enabling federated trust governance of data spaces. Organizations part of the iSHARE Network, the data owners, data providers and data consumers, all signed the same NDA & Terms of Use. Each party is validated in the registration process by iSHARE Satellites to assure trusted onboarding. Through the federated Authorization Register, Service Providers Data Owners give consent on data attributes available at Data Providers specifically to selected Data Consumers with the use of licenses.

ANALYTICSHUB

Web site: <https://cloud.google.com/bigquery/docs/analytics-hub-introduction>

Short description: AnalyticsHUB data exchange platform that enables you to share data and insights at scale across organizational boundaries with a robust security and privacy framework. With Analytics Hub, you can discover and access a data library curated by various data providers. This data library also includes Google-provided datasets.

PART II: USE CASES

PART II EXECUTIVE SUMMARY

TRUSTEE is a platform where providers can offer their data securely for third parties to perform their computations. In the following paragraphs the data providers are presented from the fields of energy, health, education, automotive and space. For each of these domains there is a description of the data provided, their usage, the communication channels and forms that are supposed to be used. Also, the document refers to data accessibility, security, and protection, as well as ethical aspects and requirements for each one of the six categories separately. The goal of the platform is not only to use a dataset individually but also to enable multidisciplinary use and exchange of data, so this part concludes with a presentation of such data, their accessibility cases and security, protection, and ethical requirements. Where required and applicable, the scope and detailed information about the use cases will be enriched in Deliverables 2.2 and 2.3.

PART II INTRODUCTION

This part of the document outlines key stakeholders and actors within each pilot of TRUSTEE. It describes and defines their unique ecosystems, data exchanges, and the lifecycle of existing data; the requirements, sensitivity, and types of data generation and exchanges each pilot has, and the laws, frameworks and other terms and conditions (at EU or local/regional level) that each pilot adheres to or implements with regards to data.

REVIEW OF REQUIREMENTS WITH PILOTS – 1ST ITERATION

Regarding the technical specification of the use cases, interviews were conducted with the partners leading the TRUSTEE Pilot Use cases, who provided details regarding the datasets that are involved in the corresponding domain of each TRUSTEE Pilot Use Case. These are detailed in the Data Management Plan (DMP D1.6⁹⁶⁶). While D1.6 describes the type, size, metadata, sensitivity, and privacy requirements of the datasets as well as the way the Pilot Use cases' datasets can be accessed, below describes the requirements, challenges, potentials, and opportunities from advancing data access for secondary use and processing. Specifications around the data environment of Pilot partners are detailed in the DMP, D1.6.

ENERGY

DATA SUMMARY

CHECKWATT are developing and marketing solutions that help Swedish households and companies become smarter and more sustainable in their energy usage. Currently, CHECKWATT are monitoring more than 4500 PV installations in real-time. CHECKWATT are managing the data collection and visualization of electricity production and consumption. Energy production is monitored with a resolution of one minute while energy consumption is monitored per minute or hour depending on available meter.

⁹⁶⁶ Spanakis, E., Sakkalis, V.-F., Forgó, N., Barud, K., Özkayagan-Prändl, H., Sailllant, C. – U., Markakis, E., Papatsaroucha, D. – H., Rosengren, P. – C., Kourtis, A., Xilouris, G. – Z., Hillman, V. – E., Nicolin, A. I. – I., Niskanen, I. – V., Martignano, M.-U., RUTTO, G., TESFAI OGBU, I.-R., Grguric, A. – E., & Lalos, A. – I. (2022). TRUSTEE Consortium (2022) D1.6 Data management plan available.

CHECKWATT is operating a reporting service of data to the Swedish Energy Agency for green energy certificates, guarantees of origin and trade of these. End-users have access to smartphone apps and web interfaces for visualization of their current and historical production and consumption. CHECKWATT are monitoring from a small household rooftop installation (5 to 20 kWp) to multi-tenant installations (20 to 100 kWp), up to some larger ground mounted solar photovoltaic (PV) parks (50 to 500 kWp).

Their systems allow them to combine data to see how much energy has been consumed compared to how much it has been produced. Additionally, the pilot has a recent development that enables flexibility services for the grid whereby they can offer grid operator reserve or power, so that one can either demand to put power into the grid or take power out of the grid. For these flexibility services, it means that they need to connect solar production, which they already have, with the information about battery storage and the consumers energy consumption. That is one of CHECKWATT's goals to develop through TRUSTEE. In a word, there are three data sets from three different sources that CHECKWATT aims to put together to analyze and provide efficient solutions for energy consumption, namely:

- Data from solar panels (solar energy)
- Data coming from battery use
- Consumers' energy consumption

Currently, it is still under investigation whether real or not real data will be used in this pilot, and it will be further discussed in future versions of this deliverable, namely D2.2 and D2.3. In case real data are used, data such as household addresses and postal codes will be removed to ensure de-identification of the individuals whose data will be used. The goal will be to get solutions for privacy and security, whereby data can be accessed and used without risks for personal privacy loss.

There are other data systems that can further be connected to a solution that can allow them to analyse these different silos of data, which will come from energy produced by wind turbines.

The expectation of data size is around 400 GB compressed and refer to data from 2020 onwards. The data comes from either one of the devices that belong to CHECKWATT, such as the meter or gateway, or from the energy provider on energy consumers' consumption data.

MAKING DATA ACCESSIBLE

Metadata will be discoverable and offered in a way that they can be harvested and indexed. There is no specific vocabulary around their data and only JSON files will be available. If there is a specific recommendation at consortium level, the pilot can accommodate it. Certain standards are implemented with regards to weather data specifically/distinctly. It would be also interesting to collaborate with ISS (TRUSTEE Space Pilot Use Case) and identify best practices, formats, or methodologies that can be followed to make data interoperable and allow for data exchange and re-use (across disciplines involved in TRUSTEE).

DATA SECURITY AND PROTECTION

In the event that real data is used in this pilot, there are various data elements that only authenticated people can have access to. The system that is currently implemented allows for the creation of a separate database that protects the real data (i.e., identifiers of households such as postcode, addresses and house numbers). There is a specific data security policy, which requires that a virtual identifier is created for each customer that will be used in all TRUSTEE spheres. Customers will only be known to CHECKWATT in a separate database. Lastly, any licensing and data sharing agreements are implemented with the energy provider and energy consumers only.

ETHICAL REQUIREMENTS

If real data are used, then CHECKWATT will require ethical and legal consent from energy consumers to use data about their energy consumption.

OTHER ISSUES

Creating a solution for accessing and use of any of the types of data is a possibility for the use case partner specifically with regards to the outcomes from the research, not the original data. In case any issues arise throughout the course of the project they will be discussed in future versions of this deliverable, namely D2.2 and D2.3, as well as investigated under WP5 activities and discussed in upcoming WP5 Deliverables.

HEALTH

DATA SUMMARY

Università Cattolica del Sacro Cuore (UCSC) is a partner providing data from the health domain for the needs of a use case pilot with the objective to identify potentiality of utilizing cross-domain data, deploying privacy-enhancing mechanisms such as HE, and identifying beneficial data processing by using data resembling health data and data from other domains.

Regarding the applicable standards, which allow to make data FAIR, the current form of the data and the datasets allow only for a partial findability and interoperability of the data. It seems that the data are only partially collected according to the standardised formats (such as the COVID Case Report Forms (CRFs) issued by the WHO), however, the standards may vary depending on the location, and the institution and its level (local/regional/country level). It is therefore important to work on the standards that would allow to unify the data and implement appropriate standards, which would facilitate the cooperation and communication between different datasets. Currently, these differences and deviations among the standards do not allow for a smooth transfer of the data among the institutions.

The issue in this setting is two-fold: the responsibility of local health authorities for the whole process of data acquisition, preservation, and communication among the different partners; and the form of communication and dataset formation.

MAKING DATA ACCESSIBLE

In the case of “COVID-19 CRF” data, these can be accessed by several different actors and processed anonymously or not. The data flow for this type of data is as follows and can be seen in Figure 56.

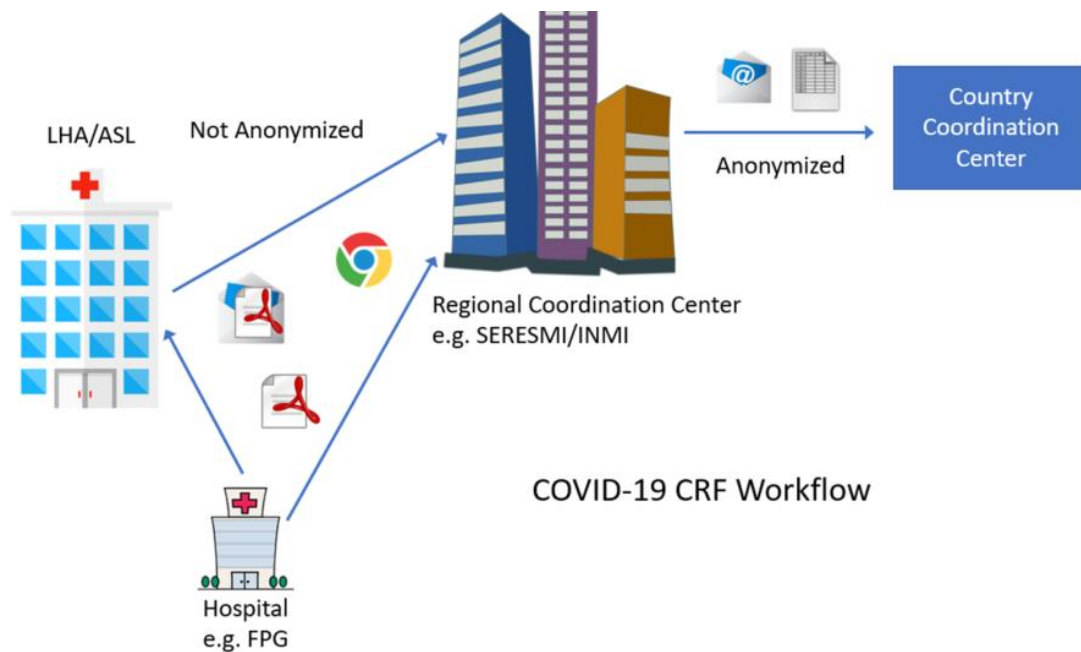


Figure 56: Stakeholders accessing data

The schema shown in the context of the use case is real. The (real) systems used by the UCSC researchers will be set available and accessed by the consortium partners, however, the data uploaded onto the database will be fictitious, as described in D1.6 – Data Management Plan available. Hence, the data schema will be real, but the data itself will not.

The forms are split into three modules:

- Initial form when the patient is admitted.
- The form highlighting the check performed daily.
- The module indicating the result.

These forms are compiled by local authority /hospital / initial examiner either electronically through a system available on a regional level, or in a PDF form. The regional coordination centre collects the information about the patient and may provide it for statistical purposes, in an anonymized format, by stripping the dataset of all personal data, leaving just the patient IDs (as required by the WHO standards).

For sharing personal data, the institutions have appropriate framework agreements regulating the data processing and sharing aspects.

The data from the database of Contact Tracing Information, defined in Figure 56, are shared among different actors in the following manner (the data sharing schema is the real one):

- Bottom-up approach: Contact tracing information is not harmonised and it is in the responsibility of local health authorities, who conduct interviews reconstructing the history and all results of the contact that are confined in this local health authority. In the case where there are several regions involved, the information goes further in an electronic (e-mail) format to the regional coordination centre. Format of the emails is not unified between the Local Health Authority (LHA)/Azienda Sanitaria Locale (ASL) (from Italian) or local health unit. The country coordination centre receives the data in the same format and records it further and coordinates further actions on the data set. This approach is illustrated in Figure 57.
- Top-down approach: The communication is conducted in an irregular form via e-mail and forwarded to the patient/hospital/GP over the phone. The issue appears when the number of infections is higher, the traffic is high, as the contact tracing is done manually, and it is very limited due to the processing capability at the different coordination centres. This approach is showcased in Figure 58.



Figure 57: Bottom-up approach contact tracing

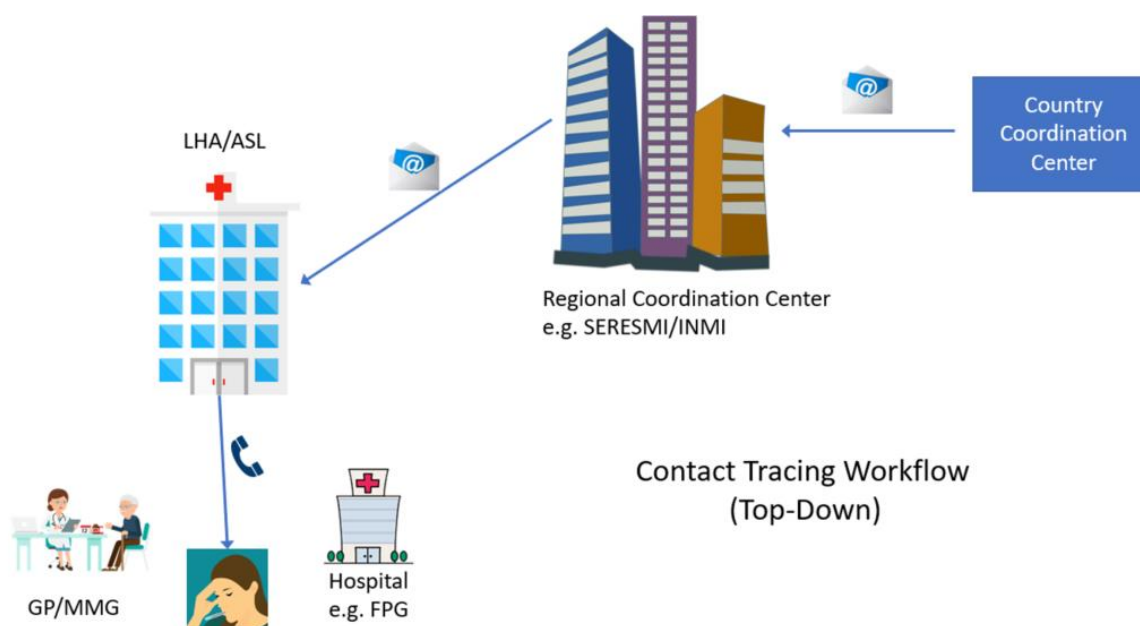


Figure 58: Top-down approach contact tracing

DATA SECURITY AND PROTECTION

UCSC requires enhanced data privacy and security of its data and is interested in assessing how the adoption of HE could enhance data privacy and security. Typically, in the real scenarios, health data adopts WHO standards for data de-identification. That is, patients' names are not used, only their assigned IDs. Each hospital which is registered at WHO has an ID assigned and assigns/allocates a patient ID subsequently. The de-identification is done at a regional level before the information from healthcare forms is further provided to the country coordination centre. However, issues persist. For instance, local implementation of the CRFs may deviate from that and contain directly identifying personal data of the patient. ID assignment is the only anonymization technique applied by the hospital at the hospital level. Additionally, the Contact Tracing information (since UCSC is using data from Covid contact tracing cases) contains in a real scenario substantial personal data, whose anonymity cannot be guaranteed entirely. For that, UCSC aims to examine how HE can solve any of the privacy concerns relating to personal and sensitive patient data.

ETHICAL REQUIREMENTS

Medical health data and data concerning health require a higher level of protection and security, as their processing may pose a high risk to the rights and freedoms of a data subject. The level of protection and number of ethical requirements depend on factors such as whether the data have been pseudonymised or anonymized.

OTHER ISSUES

With regards to multi-disciplinary data exchange, the pilot intends to collaborate with the TRUSTEE Education Pilot Use case. In case any issues arise throughout the course of the project they will be discussed in future versions of this deliverable, namely D2.2 and D2.3, as well as investigated under WP5 activities and discussed in upcoming WP5 Deliverables.

EDUCATION

DATA SUMMARY

A school environment comprises a Student Information Management (SIM) system. The school system connects with an Education Technology provider (Edtech), a third-party data management system that allows partial access to data of specific students, such as 6th grade movers (primary level) to 7th grade (secondary level).

A platform provider such as Google Classroom or Microsoft Teams manages pedagogic operations such as homework and instruction material, and communication between the student and the teacher. The same platforms can have access to other student, teacher, or school information such as students' in-house or out-of-school communications, IP address, type of devices and networks, and in-app behaviour (and similar granular metadata).

A third data system connects to the school, which provides roster, attendance, conduct, student information (including address, health records, age, name, and previous school records), achievements, timetable, any documents related to the student that the school already has on record, prep (homework), and emails from schools and teachers.

Figure 59 and Figure 60 below showcase data flows in and out of different systems including SIM and Edtech providers.

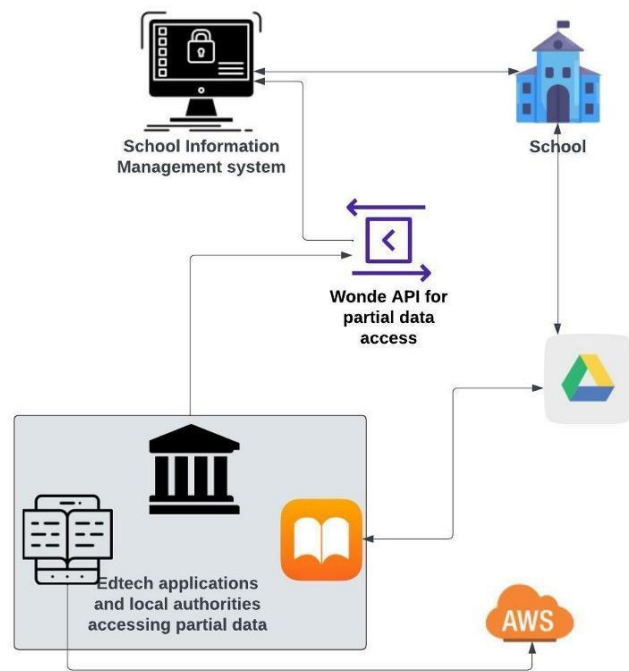


Figure 59: Simple stakeholder environment in an educational setting

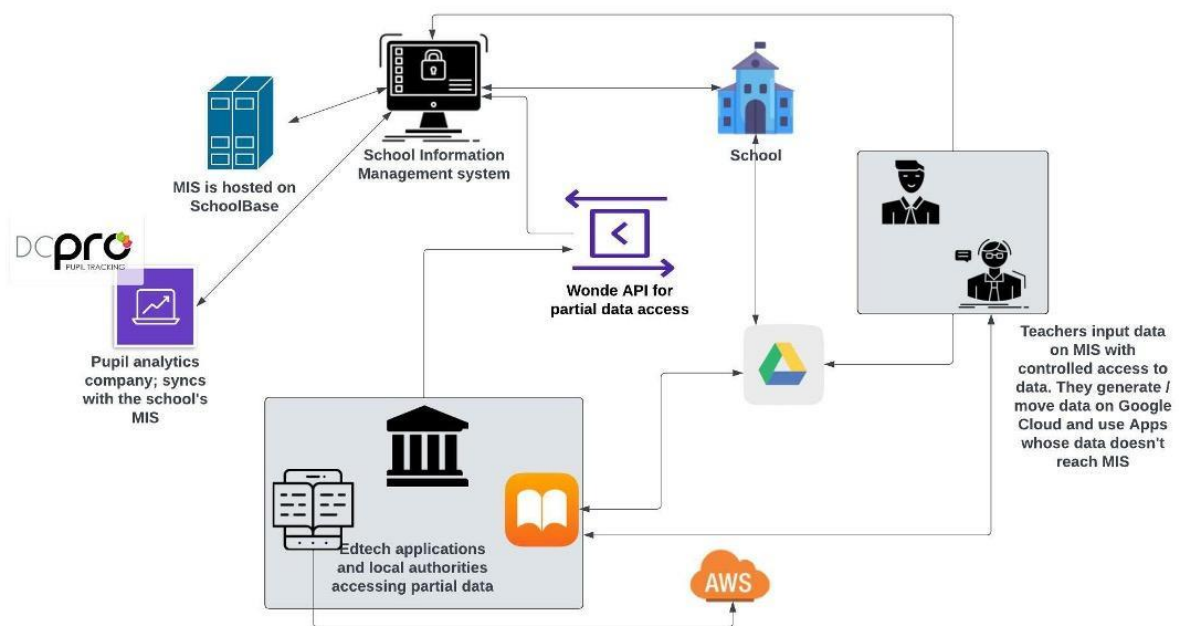


Figure 60: Complex stakeholder environment in an educational setting

Etoile Partners (EPL) will use data from one school in England along with their third-party provider, who gives partial access to 6-into-7 data of students, that is data transferred from primary to secondary schools when children from grade 6 move into secondary schools starting with grade 7. This transfer is particularly sensitive because of the risks involved in how data is processed and exchanged in the transfer.

Such third-party providers may also have view and modification rights to students' documentation, such as communication, calendar events, and assignments/homework. Therefore, EPL's objectives with TRUSTEE is three-fold:

- First, this third-party access and view of student behaviour is invasive, and a solution will be sought that can preserve the privacy of students from such platforms, while allowing student data to be used for beneficial purposes.
- Second, the exchange of data should be done in an accountable and privacy-preserving way.
- Third, the optimisation of data sharing and use across the different systems should be achieved, such as between SIMs and third-party Edtech providers or between TRUSTEE pilots, e.g., the Energy Pilot Use case.

There is no specific data size, since only the data schema and ontologies will be provided. The files will be in JSON and XLS format. Synthetic data will have to be generated to populate the files. There are various stakeholders that will be potentially keen to access (or potentially provide access to) such type of data, namely: education departments/ministries, research communities (often near impossible due to high risk ethically), NGOs and other agencies that deal with children (immigrants, special education, fostered, under protection, and other conditions). The use of these data is therefore extremely sensitive and therefore security and privacy preservation are the priority.

MAKING DATA ACCESSIBLE

There will be partial data accessibility. At first, the data schema will only contain synthetic/fictitious data. Then, the consortium must agree on which partners will need to access this data and, therefore, establish a data access control.

Data anonymisation is essential, even though data will be synthetically populated. This will allow for full assessment of the TRUSTEE platform if individual privacy is preserved.

DATA SECURITY AND PROTECTION

The UK-GDPR and the EU GDPR will be followed as the main legal framework and procedures for data security and protection. Additionally, certain vocabularies and standards are openly available (e.g., SIF, CEDS and similar schemas). Vocabularies and methodologies of school level are also available, however, not all of them are fully implemented. At Edtech providers' level it is yet to be identified if there are specific data standards used.

The data quality assurance will come from the standards that public educational systems use and then at Edtech (private industry) level, there is a need to set specific standards that companies (should) meet but that will be clarified later and further discussed in future version of this deliverable, namely D2.2 and D2.3, as well as upcoming deliverables of WP5. Possibly, such companies are also using quality assurance like CE and CE+ or ISO standards.

Plugging the multi-modal data in a federated framework to see how this would improve and assure privacy is one objective for EPL within the educational context. Dealing with data leakage problems, dealing with security risks, partial identification of data, data transfers, and other similar risks are key aspects that need to be considered when developing the TRUSTEE Platform. As a solution, it will be important to see on what part of education data transactions HE can be implemented and what it will really protect. Lastly, what types of control and at what stage of data use, exchange or computation must be implemented and on which stakeholder (Edtech provider, school, teacher - all of them as a condition).

EPL, according to the information provided in the D1.6 Data Management Plan, will provide fictitious data for the purposes of the pilot.

ETHICAL REQUIREMENTS

For the time being no specific ethical agreements and Data Management Plans (for the data schemas used to develop the fictitious database) will be required, however as the project advances this section will be revisited for further considerations and will be further discussed in future version of this deliverable, namely D2.2 and D2.3, as well as upcoming deliverables of WP5.

OTHER ISSUES

EPL has developed a legal, cybersecurity, and ethical framework and procedures that assess the education stakeholders. This is one framework that can be explored. It will be partially made available for consortium partners to review. In case any issues arise throughout the course of the project they will be discussed in future versions of this deliverable, namely D2.2 and D2.3, as well as investigated under WP5 activities and discussed in upcoming WP5 Deliverables.

AUTOMOTIVE

DATA SUMMARY

Panasonic Automotive Systems-Europe (PASEU) conducts Research and Development (R&D) on a multitude of areas on automated driving, ranging from scene understanding and perception to co-operative awareness and path planning. PASEU will provide the vehicles for testing in the pilot and the on-board sensors for scene understanding and vehicle control. Here it must be noted that as per international strategies (e.g., the US National Car Assessment Program [USNCAP] and the EU NCAP), PASEU provides data that are captured during the Pilot Partner's private testing ground and scene set-up. That is, dummies for cars and pedestrians are created to simulate scenes for automotive testing (no real vehicles or individuals are used), which subsequently generate data. The data PASEU provides on TRUSTEE mainly concern car features, rather than personal data such as drivers' data (i.e., licence plates, faces of the drivers) or driver assessments. Secondly, data manipulation will be used with both real data (i.e., data captured by sensors) and synthetic data. Though trials are planned to start with synthetic data, the aim is to enhance the solution by utilising real data. For the usage of real data, a non-disclosure agreement will follow.

No personal data will be disclosed. Any personal data PASEU shares will undergo anonymization on the Pilot Use case leader's servers.

As detailed in D1.6 – Data Management Plan available , PASEU's goal is to enhance the safety indexes for autonomous driving in the presence of sensor and data sources by assessing drivers' behaviour and Human Machine Interface (HMI) evolution. This can be achieved through development of a benchmark for validating the parameters of driver's behaviour in association to Autonomous Driving (AD) functions and use of the output as informed experimental feedback for evolving the HMI. The pilot use case aspires to metadata on the conclusions drawn from the data analysis within the pilot, to be useful to the automotive and scientific community.

As per D1.6, PASEU provides multimodal data. Data is generated via simulators (i.e., generated via open-source tools; from Panasonic's own datasets; and from open-source datasets).

Similar to the Health pilot use case, the Automotive pilot use case aims to explore how data privacy can be assured and improved via plugging multimodal data in a federated framework.

MAKING DATA ACCESSIBLE

A two-tiered approach is needed where the data coming from the simulator is made openly available through a free and standardised protocol for research purposes. For the access to data coming from real vehicles, an access protocol is also needed. A non-disclosure agreement is also required by PASEU as a rule to grant access to data, during and after the project. No issues are expected with respect to the use of data for research purposes, but the publication of datasets may be restricted. Access to data (a read-only access) will be granted only to those partners who will need to use the data.

DATA SECURITY AND PROTECTION

PASEU will put in place an identification process to ensure the security of data and will grant access to data for specific partners only. As a security measure, PASEU implements a two-stage identification process: first, access to data is required after which data is encrypted, and then a password is required to access the data. For data storage (i.e., whether data can be stored in a machine connected to the internet or not) and data access, alongside the requirements under the existing applicable legislation, additional requirements may be needed.

ETHICAL REQUIREMENTS

Specific ethical requirements are envisioned during secondary data use regarding the automotive domain. While PASEU does not intend to make personal and sensitive data available on TRUSTEE, specific scenarios were considered sensitive whereby other data may become involved in combination with automotive data. In a car accident, data collection at personal level may not be an issue when it's a matter of life and death. However, if insurance companies also want to access this data, then it may become a problem for individuals. The same thing can be at a social and personal level (e.g., if a person's infidelity is revealed because of the car accident). To this end, PASEU recognises the socio-ethical requirements that emerge and aims to abide by those. More details on socio-ethical requirements are provided in the next chapter: Part III: User-centred, Legal, and Socio-ethical requirements.

OTHER ISSUES

On a corporate level, PASEU has different data management guidelines and research management policies according to different types of projects they participate in, depending on the nature of the project (i.e., project purely on research and development). PASEU adheres to global ISO standards in relation to the structural integrity of data as well as to standards regarding data protection prescribed under the GDPR. In case any additional issues arise throughout the course of the project they will be discussed in future versions of this deliverable, namely D2.2 and D2.3, as well as investigated under WP5 activities and discussed in upcoming WP5 Deliverables.

SPACE

DATA SUMMARY

Institutul de Stiinte Spatiale (ISS) provides access to large and diverse public data, which comes from space missions, coordinated either by the National Aeronautics and Space Administration (NASA), or by the European Space Agency (ESA). The data includes published data from NASA on seismic events on Moon and Mars as well as data on the position of craters and the Moon and high-resolution images of the Moon.

It is foreseen that there is no personal data, which will be made available on TRUSTEE, while real data provided by ISS are without intellectual property rights, privacy, or confidentiality safety issues and without significant security constraints. Additionally, the shared data is already published by NASA.

ISS's data on seismic activity and activity on Venus, Moon or Mars, and solar flares provide an opportunity to combine such data with data from the solar panels of the Energy Pilot Use case to examine possible secondary data use and opportunities for discovery.

ISS will be using already existing data and aims to input their data into the sandbox for the consortium. The data can be repurposed because they have no security clearance and can be put in a place for third parties to have unlimited access to them. Because the data does not belong to ISS, there can be no special identifiers assigned to them. There will not be any specific embargo applied to the data (because they have already been published by NASA and similar agencies).

Furthermore, there are no specific data ontologies and vocabularies because the sector is so small that there tends to be community-based type of vocabularies, ontologies, and protocols. Additionally, the data are simply structured. There are international standards applied to them but again these are more community-based standards, which are commonly known in the sector (e.g., a standard for earthquake data which follows central Pacific Standard).

The only issue with community-endorsed standard is that it may endorse certain formats that belong only to the community, and they are often informally known. There is no such formal protocol that outlines these community-based standards.

MAKING DATA ACCESSIBLE

The data will remain available for as long as they remain publicly available according to its publisher – NASA.

DATA SECURITY AND PROTECTION

ISS uses clear security protocols; however, the nature of the data is not sensitive or personal and there are no internal specific guidelines and reviews or specific conditions for which the data can be processed. Nevertheless, the combination of space data with energy data (see next sub-section) may change these conditions. This has been also indicated in the D1.6 - Data Management Plan available.

ETHICAL REQUIREMENTS

In certain circumstances, secondary data use and manipulation (between space and energy data for instance), some ethical requirements may have to be observed since the involvement of mapping the energy consumption with space data may lead to sensitivity around identifying personal data (e.g., location, household names and addresses).

OTHER ISSUES

No other issues are identified in this stage of the project regarding the data provided by ISS for the Space Pilot Use case. In case any issues arise throughout the course of the project they will be discussed in future versions of this deliverable, namely D2.2 and D2.3, as well as investigated under WP5 activities and discussed in upcoming WP5 Deliverables.

MULTI-DISCIPLINARY DATA EXCHANGE

DATA SUMMARY

Teknologian Tutkimuskeskus VTT (VTT) together with Zortenet Idiotiki Kefalaioxiki Etaireia (ZORTENET), and Hewlett Packard Italiana (HPE) provide the Trusted Multi-disciplinary Data Exchange Pilot Use case with the aim to bring forward different data models and ontologies from the telco and I4.0 domain. The objective is to build an IDSA (“International Data Spaces Association”) environment with IDS connectors to ensure trusted data exchange. Apart from an existing small scale 5G testbed, ZORTENET also aims to set up a small-scale supply chain within their lab, to recreate, together with VTT, a full IDSA environment with the connectors and end points for data exchange between different modules. The goal is to identify how the different pilot use case partners can gain benefits by integrating their data onto TRUSTEE – in other words, how to combine energy data with automotive data, space, health data and so on.

MAKING DATA ACCESSIBLE

VTT’s objectives, as showcased in Figure 61, are to replicate (partially) VTT’s Minimum Viable Dataspace (MVDS) in ZORTENET’s premises; enable IDS Connectors between 5G and the supply chain; extend to multiple setups; and exchange information between different supply chains. This Pilot Use case provides no data of its own.

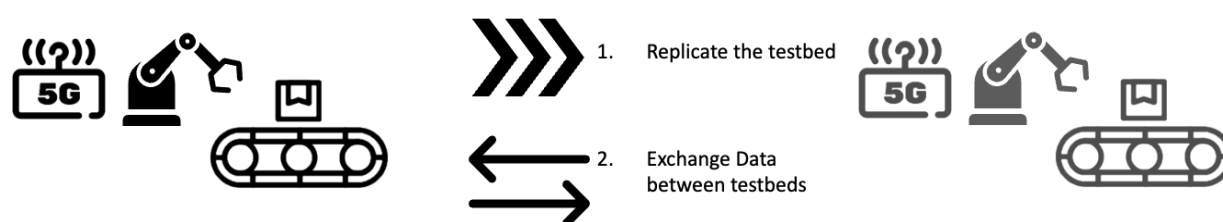


Figure 61: VTT objectives for data exchanges

DATA SECURITY AND PROTECTION

ZORTENET does not store any personal data. They store their data on systems with backup configurations and grant access to their infrastructure through a secure encrypted connection via VPN, with a predefined admin as well as control rights assigned to different users. There are no specific quality assurance processes in place. On an institutional level, the pilot use case does not engage in processing of any personal data (including sensitive data) nor storing any user data. Considering the aim of this use case being the exploration of multi-disciplinary data exchange and the end goal of having data from all other pilot use cases, ZORTENET will follow instructions of other partners in processing data when required. Following the identified sensitivities around secondary data use and AI deployment, ZORTENET will adhere to the legal requirements.

ETHICAL REQUIREMENTS

ZORTENET does not require any ethics approval to share data. However, following the identified sensitivities around secondary data use and AI deployment, ZORTENET will adhere to the identified socio-ethical requirements, which will be further described in the next chapter: Part III: User-centred, Legal, and Socio-ethical requirements

OTHER ISSUES

No other issues are identified at this stage regarding this Pilot Use case, since there will not be any data provided by VTT, ZORTE, and HPE. Any issues that might emerge regarding the multi-disciplinary data exchange that will be performed for the scope of this Pilot Use Case will be discussed in future versions of this deliverable, namely D2.2, and D.2.3. as well as investigated under WP5 activities and discussed in upcoming WP5 Deliverables.

TRUSTEE USE CASES CONCLUSION REMARKS

This part outlined preliminary background research in understanding the different data environments and domains which are taking part in the TRUSTEE project. As a platform, TRUSTEE aims to allow data providers, as the Use cases described above, to offer their data securely for external stakeholders such as research community, peers, and others to perform various investigations and computations. Each of the introduced domains in this document has allowed the TRUSTEE consortium to delve further into the various scenarios, define specific stakeholders – personas, users, and third parties who may be impacted in some way by the new possibilities of such data exchange – and investigate and define concrete although preliminary User-centred Legal, Socio-ethical, and Technical requirements that must be addressed by the TRUSTEE community of developers, data providers, AI model providers, and third party end-users or stakeholders of possible secondary data use and AI manipulation. Following, Part III: User-centred, Legal, and Socio-ethical requirements as well as Part IV: Technical Requirements discuss in depth these critical points and outline in detail the User-centred, Legal, Socio-ethical, and Technical requirements that need to be considered by a platform such as TRUSTEE.

PART III EXECUTIVE SUMMARY

This chapter describes and analyses the methodology that was followed for the extraction of the User-centred, Legal, and Socio-Ethical requirements for the TRUSTEE Platform. The first section of this chapter provides a brief description of UX Research and the importance of adopting it in a project that needs to identify end-users' needs and objectives; the second section describes the methodology that was used for the extraction of the aforementioned types of requirements; and the third section comprises the outcomes and analysis of the applied methodology, which resulted in the formulation of the requirements.

PART III INTRODUCTION

The engineering of the requirements presented in this chapter was based on state-of-the-art research and analysis of the currently binding relevant legal acts, as well as pending legislation-procedures regarding Legal and Socio-ethical requirements, already presented in Part I: State of the Art, and on UX research regarding User-centred requirements. Nevertheless, UX research on end-user's needs and objectives led to a greater comprehension of end-user perspectives on the Legal and Socio-ethical requirements as well for a Platform such as TRUSTEE.

USER EXPERIENCE RESEARCH: WHY AND HOW TO CONDUCT IT

UX is associated with technological products and their usability, design, and user experience. It is related to Human Computer Interaction (HCI) and deals with the outcome of a user's contact with a technological solution, product, or service. UX considers all aspects of the product from the aesthetic to the interactive and the context in which it is used, consequently linking the product with a user's needs and goals. Users are set to the centre of a product's development process, so it is important to include their perspective in product design⁹⁶⁷.

User-centred Design encloses users in every aspect of a product's lifecycle, focusing on their needs and requirements and aims to a usable product that is user-driven. Product designers spot and acknowledge users' needs by conducting user research, namely applying methods and tools to identify user requirements for the product in order to develop the solutions. These methods include interviews, surveys, focus groups and other techniques, each one offering a different perspective of the end-users' requirements. Interviews have the benefit of getting close with the user asking questions, extracting detailed information about their thoughts, feelings, and desires regarding the product. They may be time consuming, but the level of information and specifications assists extremely to the development procedure. On the other hand, surveys have a structured format, and they can collect information from large groups of people, thus providing quantitative information regarding user requirements, in contrast with the qualitative type of information that interviews provide. Focus groups are a gathering of approximately ten end-users that may answer questions or respond to a product prototype, relate to specific enquires around why certain issues or challenges emerge, and allow for more open-ended and deeper discussions around specific contexts. Their thoughts and observations can lead to ideas and suggestions that will evolve the product and take the whole creativity process to the next level.

⁹⁶⁷ Hassenzahl, M. & Tractinsky, N. User experience - A research agenda. *Behav. Inf. Technol.* **25**, 91–97 (2006).

All these methods are communication bridges between developers and end-users, so that the first can form their plan and goals regarding the product development, while the latter will end up with a usable and quality product. This information is collected, organized, and used to form Personas and User Stories ⁹⁶⁸.

Personas are a user-modelling technique that contributes to product design of end-user products. Personas are fictional characters based on real data gathered from interviews, demographics or information from marketing teams or customer-related services and surveys. Their creation includes an image or illustration, personal characteristics and information, thoughts, description, goals, and needs. Personas succeed in representing real users and Use case Scenarios, identifying problems and user-centred requirements, giving the designers and developers the chance to empathize with them, understand their needs and come up with solutions to those desires and problems ⁹⁶⁹. They identify and illustrate specific scenarios and highlight connections among different user types. User Stories are the stories that relate to Personas and aim to picture the Persona while they are working or performing a task as well as their behaviour and thoughts regarding that context. This way, the project developers may identify the technical product requirements while they highlight the requirements' priority stage. Additionally, project members perceive the same conceptual model assisting in the team's cooperation and can easily remember and revisit the user's perspective⁹⁷⁰.

Using Personas in product design benefits a team's efficiency, productivity, and creativity. The designers and developers relate to the fictional characters, understand and identify their needs, and therefore are able to come up with solutions and product aspects that ensure functionality and a pleasant experience⁹⁷¹.

In a similar manner, UX in data handling refers to the user - data interaction which is related with a variety of actions like data insertion, collection, organization, quality, accuracy and understanding. Moreover, it is important for the user to access interfaces and tools that assist in data analysis, inform decision making, and facilitate and advance work. This comes with challenges like data complexity, formation, and structure, as well as diversity of data sources and specialization requirements in working with data.

The use of Personas can identify and specify all these challenges and requirements, while also providing use stories, forming a clear and descriptive view of the project. Project developers are able to identify possible limitations in tools and technologies for data handling and analysis or privacy, cybersecurity and socio-ethical issues regarding use, assessment, and exploitation of data for secondary purposes.

METHODOLOGY

This section thoroughly describes the methodology, towards UX research, followed in TRUSTEE for extracting User-centred Requirements. The methodology of the state-of-the-art research analysis followed for identifying Legal and Socio-ethical requirements stems from the legal framework and procedures presented in Part I: State of the Art. Nonetheless, and since UX Research assisted in unravelling the perspective of end-users towards Legal and

⁹⁶⁸ Courage, C. & Baxter, K. *Understanding your users: A practical guide to user requirements methods, tools, and techniques*. (Gulf Professional Publishing, 2005).

⁹⁶⁹ Aquino, P. T. & Filgueiras, L. V. L. User modeling with personas. *ACM Int. Conf. Proceeding Ser.* **124**, 277–282 (2005).

⁹⁷⁰ Schneidewind, L. *et al.* How personas support requirements engineering. *2012 1st Int. Work. Usability Access. Focus. Requir. Eng. UsARE 2012 - Proc.* 1–5 (2012) doi:10.1109/UsARE.2012.6226786.

⁹⁷¹ Long, F. The effectiveness of using personas in product design. *Irish Ergon. Rev. Proc. IES Conf.* 2009 1–9 (2009).

Socio-ethical requirements for a platform such as TRUSTEE, this section also discusses the role of UX Research in the approach towards identifying the Legal and Socio-ethical requirements relevant in TRUSTEE.

UX RESEARCH FOR USER-CENTERED REQUIREMENTS EXTRACTION

Based on the literature presented above and for ensuring that the TRUSTEE Platform will be tailored to end-user needs, the methodology that was followed for extracting the User-centred requirements was focused on UX research and consisted of the following steps:

- One-to-one interviews were conducted with the partners leading the Pilot Use Cases of TRUSTEE, which focused on identifying the initial technical specifications of the Use Cases, recognizing the Pilot Leaders' needs from a solution such as TRUSTEE, and delving further into the definition of the TRUSTEE Pilot Use cases, as presented in Part II: Use Cases.
- Surveys were developed by gathering a large pool of questions that aimed in diving deeper in the goals and frustrations of end-users that the TRUSTEE platform will eventually serve and determining the TRUSTEE Personas and User-centred requirements.
- A Co-design Workshop was organized for getting together with potential TRUSTEE end-users and stakeholders, where participants were distributed into focus-groups with the aim to discuss, validate, and enhance TRUSTEE Personas and User-centred requirements.

INTERVIEWS

The interviews conducted by EPL, UNIVIE, and HMU with all TRUSTEE Pilot Use case Leaders achieved two objectives. First, they allowed for EPL and UNIVIE to grasp more in-depth the environments of each Pilot Use case, and identify their challenges, needs, requirements, and understanding of legal and socio-ethical conditions in relation to data management and use, as well as their expectations from and ambitions for the TRUSTEE project. These initial findings, outlined in Part II: Use Cases, further laid an initial understanding of the Pilot Use case environments and allowed us to set up the Co-design Workshop 1.0 and widen the scope of discussion around requirements, challenges, and ways forward for the development of the TRUSTEE project. The second objective of interviews was to implement the first overview of the initial technical specifications of the Use cases, the types of data that will be included in each of the Use cases and with which TRUSTEE solution would operate, and the prospects for how the data will be made FAIR within the project. Moreover, the partners leading the Pilot Use cases shared their initial data security and data protection concerns relevant for their respective Use case in the context of using TRUSTEE for making data available for further data computations and manipulations, as well as for AI application. Based on that, UNIVIE could initially assess the legal regimes that would need to be considered as part of a preliminary legal framework and procedures for the TRUSTEE project and TRUSTEE solution while EPL could initially assess the socio-ethical requirements around secondary data use and AI data manipulation. The analysis of the legal and socio-ethical matters has been respectively conducted as part of T1.5 and T2.2 efforts and provided in this Deliverable for the TRUSTEE solution in Part I: State of the Art and in Part II: Use Cases, and in the D1.6 Data Management Plan available for the TRUSTEE project. Furthermore, based on the process described:

- UNIVIE extracted from relevant legal acts, relevant standards, and guidelines a list of Legal requirements for TRUSTEE System Developers as well as for TRUSTEE End-users (representing Personas defined in this Deliverable further below in this chapter), which were further enhanced through the Surveys and the Co-design Workshop 1.0.

- In a similar manner, EPL extracted a list of Socio-ethical requirements for TRUSTEE System Developers and TRUSTEE End-users as well, which were also complemented by the insights gathered through the Surveys and the Co-design Workshop 1.0.
- HMU was able to identify initial User-centred requirements, mainly referring to data provision and secondary data use, which assisted in developing the Surveys and designing the respective set of questions that were discussed during the Co-design Workshop 1.0, both with the aim to identify the TRUSTEE Personas and extract the complete list of User-centred requirements.

SURVEYS

A TRUSTEE survey was created for enhancing the output of the initial interviews and for assisting in the realization of the objectives of T2.3, regarding the definition of the Use cases and the extraction of User-centred requirements, as well as of T2.2, whose objectives focused on developing the legal and socio-ethical framework and procedures for data manipulation, secondary data use, and the deployment of AI functionalities within TRUSTEE. Moreover, the approach of surveys was selected as a second step as it offers the possibility of gathering more and diverse end-user responses and extend the pool of end-users, compared to individual interviews.

Regarding the objectives of T2.3 and the definition of the TRUSTEE Use cases, the primary goal of the surveys was to identify the user role groups of TRUSTEE and, thus, develop the TRUSTEE Personas and identify their challenges and goals. Based on the Personas that were developed, a User story extraction process followed, through which several User stories were extracted for each Persona based on their goals and frustrations. These User stories then led to the extraction of the User-centred requirements for the TRUSTEE Platform.

With regards to the T1.5 and T2.2, the aim of the questions included in the Survey was to collect further details about the character of the data processed by the respondents, as well as the relevance of data protection and data security requirements and the level of awareness of these in the data processing activities conducted and/or planned to be performed by the respondents (i.e., data manipulation, secondary use of data, deployment of AI functionalities). This served for further development of the list of rules, terms and conditions that reflect legal, ethical, cybersecurity and other requirements binding in the process of TRUSTEE development and its use by the TRUSTEE End-users, defined thanks to the Survey, and definition of an initial common comprehensive framework for the TRUSTEE Platform.

For realizing this set of diverse objectives, the survey questions were formulated by HMU, EPL, and UNIVIE, who closely collaborated to design, develop, and distribute the survey across the partners that lead the TRUSTEE Pilot Use cases. The Pilot Leaders were requested to complete the survey, as well as circulate it among the Members of their institutions for gathering more end-user responses.

Two different versions of the survey were created: a version that was filled out by the Pilot Leaders and a shorter version that was distributed to the Members of the Pilot Leaders' institutions. This approach was followed to aid understanding and avoid overwhelming the Members of the Pilot Leaders' institutions that were assumed to differ in expertise and in familiarity with data spaces, as well as with the solution proposed by TRUSTEE, compared to Pilot Leaders themselves.

The information about processing of the data provided by each of the Pilot Use case Leaders, as well as Pilot Use case Members, was created by UNIVIE and distributed to all respondents, including Pilot Leaders and Pilot Members (Annex 1: Information Concerning the Processing of the Data Shared in the TRUSTEE Survey).

Both survey versions are described in more detail in the following sub-sections.

PILOT LEADER SURVEY VERSION

The Pilot Leader Survey version was distributed to all TRUSTEE Pilot Use case Leaders and included 11 sections with a total of 48 questions (Annex 2: Surveys - 2.1 Pilot Leader) relating to:

- the profile of the TRUSTEE end-user, including job role, age group, years of experience, and work routine
- information about data use activities of the end-user (e.g., data entry, manipulation, collection, modification, querying, sharing, analysis, and data computations) and the amount of time spent on such activities
- goals, potential benefits, and obstacles in sharing and accessing data for secondary use and the algorithms (AI and ML) that might be employed for aiding these processes
- legal and ethical considerations regarding data use and sharing (which mostly served the goals of T2.2)
- the expectations from the TRUSTEE solution relating to foreseen benefits and requirements, as well as the identification of potential stakeholders and end-users of a such platform

PILOT MEMBER SURVEY VERSION

A slightly different version of the survey was disseminated to the Members of the Pilot Leaders' institutions considering that these Members may not be familiar with the TRUSTEE solution or with technical and legal details discussed in the Pilot Leaders Survey version. Thus, several questions from the question set relating to the use of real data, data analysis, and legal and ethical considerations were excluded from the Pilot Member Survey version. An initial distribution of this version highlighted further complex questions that the Pilot Members had difficulty to answer to, and, therefore, more questions were excluded from the initial version of the Pilot Member Survey to aid understanding and, eventually, gather more responses.

The final version of the Pilot Member Survey was sent to the Pilot Leaders who were asked to distribute it to the Members of their institutions. This version included 10 sections with a total of 32 out of the 48 questions that were included in the Pilot Leader Survey version Annex 2: Surveys - 2.2 Pilot Member).

PERSONAS AND USER STORIES

The survey responses were gathered and grouped according to the TRUSTEE Pilot Use case domain they corresponded to. For each domain, the survey responses were carefully studied and important information regarding end-users were extracted, which led to the identification of certain user role groups that guided the development of the TRUSTEE Personas.

Each Persona included goals, frustrations, and ambitions for privacy preserving data computations as well as secondary use of data. This information and insight were used to formulate the User stories of each Persona. While a Persona can highlight the characteristics, goals, and needs of a user, User stories aim for investigating the tasks that these users would like to achieve by using the TRUSTEE Platform. Following the structure of: "As a <User Persona> I want to <Goal> so that <Benefit>", several User stories were extracted from each Persona for each Pilot Use case domain of the TRUSTEE project, namely: Energy, Healthcare, Education, Automotive, Space, and Multi-disciplinary Data exchange.

USER-CENTRED REQUIREMENTS EXTRACTION

The next step in the UX Research approach that was followed involved using the developed User stories to extract the TRUSTEE User-centred requirements. While User stories represent the goals that a user wants to achieve, the User-centred requirements describe the platform's functionalities that revolve around the users' needs.

Each User story may produce one or more User-centred requirements. After gathering all the User-centred requirements produced by the User stories, extracted from the TRUSTEE Personas of each Pilot Use case domain, any duplicate records were removed or merged. In parallel to UX research work was performed under T2.4 regarding the Architecture Building Blocks (ABB) that could eventually constitute the TRUSTEE Platform, namely

Homomorphic Enabled Data Fusion (HEDF), AI Models as a Service (AIMaaS), Trustworthy AI Support Design Framework (TAI-SDF), Knowledge Repository (KR), Data Privacy Impact Assessment (DPIA), Dashboard, SSI Homomorphic Capable Framework (SSI-HE), One-Stop-Shop (OneSS), Data Lake (DL), Security and Trust Manager (STM), Authentication Manager (AM), and Accountable Transactions Recorder (ATR). These ABBs are described in more detail further below in this document in Part V: TRUSTEE Architecture. During the User-centred requirements definition, a preliminary investigation was performed relating to the ABBs that may fulfil each User-centred requirement., which later assisted in translating User-centred into Technical requirements. Finally, each User-centred requirement was tagged with a specific priority based on the respective User story that produced the requirement, as follows: Mandatory, Desirable, Optional, Possible Future Enhancement.

CO-DESIGN WORKSHOP AND FOCUS GROUPS

A Co-design Workshop was scheduled with the primary goal to enhance the TRUSTEE Personas and validate the extracted User-centred requirements, among several other objectives that are presented further below. This first TRUSTEE Workshop, namely Workshop 1.0 “Co-designing Data Spaces: Empowering Usability” and facing wider publics, focused on the opportunities and challenges of developing open data markets in line with the EU’s Data Economy strategy by: promoting and showcasing TRUSTEE and its unique contribution within the open data economy; opening a dialogue between TRUSTEE and key stakeholders to identify needs, requirements, opportunities and challenges from open data, data manipulation, AI, and the TRUSTEE Platform; building capacity and knowledge exchange opportunities relating to common data spaces and the data economy with relevant stakeholders and user entities. The objectives of this workshop were as follow:

- a) Present TRUSTEE Personas and User-centred requirements for cross-validation and co-design (core objective)
- b) Elicit proposals for Use case implementations on how multidisciplinary data sharing could be beneficial.
- c) Learn about the operating models and capabilities of data utilization.
- d) Perform a secondary data use exercise by using questions from the TRUSTEE Survey
- e) Discuss risks and opportunities attached to data spaces regarding the current state-of-the-art.
- f) Obtain additional ideas about how to address various Legal and Socio-ethical requirements from the current and upcoming legislation in the data spaces environment.
- g) Exchange experiences and knowledge capacity with others involved in the field of data spaces in respect to sustainability issues concerning the maintenance of such spaces outside research projects beyond their lifecycle.
- h) Discuss the challenges and risks of creating and orchestrating open and dynamic datasets for continual Federated Learning tasks.
- i) Promote and showcase TRUSTEE’s solutions to relevant stakeholders.

The workshop was structured into three Sessions, where the first Session gathered all participants together, whereas the last two distributed all participants into interdisciplinary and cross-domain working groups to discuss in more depth questions pertinent to the workshop objectives. The sessions focused on the following topics:

1. Session 1: User Centred Requirements & Persona Definitions (All together)
2. Session 2: Secondary Data use (breakout Working Groups)
3. Session 3: Legal and Socio-ethical requirements (breakout Working Groups)

The 1st Session focused on the present Personas and User-centred requirements that TRUSTEE had identified with the aim to exchange knowledge and ideas with external and similar projects as well as empathize with potential end-users of the TRUSTEE Platform. Questions such as needs, challenges, and opportunities for having an open data platform such as TRUSTEE guided the collection of information and the cross-validation of User-centred requirements. For the validation to be feasible, several similar EU projects were invited to the workshop, while the pool of participants was enhanced by external attendees that were considered potential end-users of the TRUSTEE Platform.

The main discussion of the 2nd Session surrounded questions relating to multi-disciplinary and secondary use of data that is already made available on TRUSTEE Platform. Participants were asked to identify further opportunities for cross-domain data processing and computation; to identify challenges that may emerge as a result of that; and define present limitations and challenges in cross-domain data processing. Participants were asked to consider specific examples and scenarios of using data from other domains besides theirs and to identify research opportunities or end-user benefits from such cross-domain secondary use.

During the 3rd Session, participants were asked to identify basic Legal and Socio-ethical requirements, challenges, and needs foreseen by the system developers as well as potential end-users of the developed solution from projects such as TRUSTEE that would need to be complied with and implemented. The session aimed to elicit two types of information: first, to understand the level of knowledge, awareness, and skill sets that participants had when handling and processing data in their domain, but also to consider the ethical and legal challenges once they do so in combination with data from other domains; and second, to identify the new legal and socio-ethical challenges that may emerge from conversations around secondary data use or repurposing data once it is made available on platforms such as TRUSTEE.

APPROACH TO LEGAL AND SOCIO-ETHICAL REQUIREMENTS DEFINITION

While the proposed framework and procedures guiding TRUSTEE system developers and users to adhere to lawful and socio-ethical use of AI and use of the tools offered by TRUSTEE has been detailed in Part I: State of the Art, here we additionally lay out the logics that have guided the development of Legal and Socio-ethical requirements and emphasise the priority our project puts to meeting legal and ethical principles. TRUSTEE has sought to address several effective actions to achieve consumer-privacy and data-protection requirements. These span the lifecycle of data, and expand the processes of operations, infrastructure, and customer-facing practices, and are enabled by data mapping. TRUSTEE's architectural designs aim to satisfy the requirements of individuals, industry, research for federated & FAIR data access through a privacy and data protection compliant framework.

LEGAL REQUIREMENTS

The main basis for extracting the Legal requirements for different actors involved in TRUSTEE was as follows:

- Initial analysis of the legal acts potentially relevant for TRUSTEE system developers and potential end-users,
- Collection of a set of inputs provided by the Pilot Use Case Leaders, Pilot Members, and external stakeholders, in the course of the initial interviews, surveys, and a Workshop presented above,
- Continuous involvement in the information exchange shared by the Partners developing TRUSTEE architecture.

All these points supported the process of creating the initial legal framework and procedures for TRUSTEE solution– from the perspective of its developers, as well as its users, included in Part I: State of the Art - Legal Foundation on Data Privacy and Security in TRUSTEE, as well as two initial lists of Legal requirements relevant for TRUSTEE development and use, which should be considered and complied with by their addressees, indicated in the respective columns of the list included in Annex 5: Legal Requirements.

The legal framework and procedures present the analysis of relevant legal acts which are already adopted, in force and applicable, and which provide a set of responsibilities and obligations for the actors involved in the TRUSTEE lifecycle. The framework includes also selected legal acts which are, at the moment, at the proposal stage, and they will, however, play a crucial role for entities such as TRUSTEE and actors involved in the development and use of TRUSTEE. The framework and procedures aim to present a comprehensive legal landscape which has to be considered in the context of data manipulation and AI provision and use in the TRUSTEE environment. It should be stressed that these refer to the initial framework and procedures, and it will be further developed and enriched in

the Deliverable 2.2 and 2.3, according to the next developments and findings of TRUSTEE and Use case Pilots in their different phases, identified and defined in WP5.

The legal framework and procedures focus, particularly, on the following fields:

- Data protection (with the EU GDPR considered)
- Free flow of non-personal data (and the regulation applicable),
- Data governance rules, stemming from the acts such as Open Data Directive, Data Governance Act as well as proposed Data Act (which is still subject to changes and is to be adopted by the EU institutions),
- AI development, provision, and use (guided by the proposed AI Act, as well as EC Ethics guidelines for trustworthy AI),
- Cybersecurity (here the main requirements stem from NIS and NIS 2 Directives as well as EU Cybersecurity Act),
- Secure identity management – regulated in the eIDAS regulation,
- Online intermediation services and requirements from the DSA,
- Lawful blockchain application, which has to take into account data protection and privacy principles,
- Contractual arrangements at the B2B and B2C level
- Intellectual property regime, regulated by Berne Convention, the Infosoc Directive, Computer Programs Directive, Database Directive, CDSM Directive, and the Trade Secret Directive.

All above mentioned laws are referenced in Part I: State of the Art - Legal Foundation on Data Privacy and Security in TRUSTEE.

The subsections below describe the chronology of collecting the inputs and their analysis, resulting in creation of the lists of requirements included in Annex 5: Legal Requirements.

The lists are separate due to the difference between the addressees, the scope of the legal norms applicable to them, and the phrasing of the requirements, depending on the addressee type.

TRUSTEE LEGAL REQUIREMENTS FOR TRUSTEE END-USERS

The Legal requirements defined for TRUSTEE end-users (Data Provider, Consumer, Model Provider and Developer) were build based on the following actions:

- a) Initial interviews with TRUSTEE Pilot Use case representatives, being partners to the TRUSTEE consortium, to identify foreseen activities, types of data which are to be involved in the Use cases, approach to FAIR data management and expectations from the TRUSTEE solution from the perspective of support in addressing and solving legal issues related to the data exchange - data sharing, access and use - foreseen to be enabled via TRUSTEE.
- b) Mapping and initial analysis of the binding, i.e. adopted, in force and applicable as well as adopted, in force and awaiting the moment of applicability, and pending legal acts, i.e. acts at the stage of proposal discussed by the EU bodies from the domain of data protection, free flow of non-personal data, cybersecurity, AI, contract law, data governance and data sharing, laws applicable to blockchain and smart contracts, intellectual property law, included in Part I: State of the Art - Legal Foundation on Data Privacy and Security in TRUSTEE of this deliverable. Provision of additional questions in the Survey provided to Pilot Use case Leaders and Members, allowing to identify further legal issues relevant for data exchange in specific domains.
- c) Workshop 1.0, and specifically Session No 3, which focused on selected legal issues related to the data sharing aspects, allowing for further identification of further legal frameworks and procedures having significant impact on the shape of the TRUSTEE solution, which have to be respected by the TRUSTEE system developers, as well as on rights and responsibilities of actors being future users of it, i.e., Data providers, Model providers, Consumers and Developers.

TRUSTEE LEGAL REQUIREMENTS FOR TRUSTEE SYSTEM DEVELOPERS

A second list of Legal requirements focuses on the group of requirements posed towards TRUSTEE system developers and parties maintaining it, to build the system in compliance with legal and regulatory framework identified and indicated in Part I: State of the Art - Legal Foundation on Data Privacy and Security in TRUSTEE. The requirements were developed based on:

- a) Regular exchange considering conceptualisation and system architecture within WP2, taking place in the weekly meetings,
- b) Exchange of written information via e-mail and shared TRUSTEE project repository,
- c) Exchange of information in the meetings concerning each of the architecture building blocks under the umbrella of WP3 and WP4.
- d) Information shared in the preparation phase of D1.6 - Data Management Plan.
- e) The analysis of the User-centred requirements and the expectations towards TRUSTEE phrased therein.

As indicated above, having considered that some of the law texts taken into account are at the proposal or adoption phase at the EU level, or are to be transposed into the national legislation, the list of requirements may be subject to further modifications, which will be reflected in D2.2 and D2.3.

SOCIO-ETHICAL REQUIREMENTS

The Socio-ethical requirements for deploying algorithmic data manipulation on TRUSTEE are drawn from several principled frameworks and resources including the European Commission's Ethical Guidelines for Trustworthy AI⁹⁷², the HRESIA framework and several others⁹⁷³.

The methodology used to develop the socio-ethical framework and procedures comprises mapping of legal, ethical, as well as technical requirements, addressing individual Personas or Stakeholders of TRUSTEE (including TRUSTEE system developers, and TRUSTEE end-users) and underpinning several overarching ethical principles: respecting human autonomy, prevention of harm (including residual), FAIR data, explicability, and human oversight. The work output provides a practical list of requirements or toolkit (Annex 6: Socio-Ethical Requirements) to which each identified TRUSTEE Persona or Stakeholder can attend to and be guided by. The rationale goes beyond data privacy and security dimensions to include a human-rights and socio-ethical (HRESIA) logic⁹⁷⁴, which helps to deal with new and emerging issues concerning the collective and secondary use of data once data becomes available from various domains on the TRUSTEE Platform.

⁹⁷² "High-Level Expert Group on Artificial Intelligence, E.C.: Draft ethics guidelines for trustworthy ai", draft document (2018)

⁹⁷³ Siau, K., Wang, W. (2018). Building Trust in Artificial Intelligence, Machine Learning, and Robotics. *Cutter Business Technology Journal* 31(2): 47–53. Accessible from: https://www.researchgate.net/profile/Keng-Siau-2/publication/324006061_Building_Trust_in_Artificial_Intelligence_Machine_Learning_and_Robotics/links/5ab8744baca2722b97cf9d33/Building-Trust-in-Artificial-Intelligence-Machine-Learning-and-Robotics.pdf; European Agency for Fundamental Rights (FRA). (April 2017). *EU charter of fundamental rights*. Accessible from: <https://fra.europa.eu/en/eu-charter>; EC (2019). *Ethics guidelines for trustworthy AI: high-level expert group on Artificial Intelligence*. Draft report. Accessible from: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

⁹⁷⁴ See Mantelero, A. (2022). *Beyond data: Human rights and social impact assessment in AI*. Berlin, Germany: Springer. <https://doi.org/10.1007/978-94-6265-531-7>

The Socio-ethical requirements underpin Legal requirements across data privacy dimensions and across secondary data use and the deployment of AI models, thus pre-empting requirements following the forthcoming adoption of legislature such as the EU AI Act. To this end, the Socio-ethical requirements specifically address secondary use and data manipulation.

The framework and procedures address not only aspects at the level of risks of discrimination that can be associated with automation and the deployment of AI models, but also the risks of harm over other fundamental rights and freedoms as enshrined in the EU Treaties⁹⁷⁵, such as the right to integrity of individuals, to education, equal rights before the law, freedom of movement, the right to privacy and thought formation, the right of expression of assembly and freedom in the workplace⁹⁷⁶, also indicated as a relevant part of the legal framework and procedures for TRUSTEE. Because of this wider context that the TRUSTEE platform presents – initial data collected originally from within one domain is now positioned in a broader scale with data from other domains – the challenge is to respect and protect not only individuals but also the collective.

Algorithms can discriminate by clustering individuals across presumed characteristics other than traditional sociological categories and from there infer, predict, or steer decision-making which can have negative consequences to the individuals of such groups. This is a particularly salient challenge for a platform such as TRUSTEE because existing data from one domain will be added and manipulated with existing data from another domain (or domains) and therefore grouping or categorising across specific characteristics should be controlled and assessed beyond the provisions of existing data privacy laws. Therefore, the proposed Socio-ethical requirements aim to encompass all these challenges and see beyond PIAs.

The framework and procedures aim to guide multiple Personas, both technical and non-technical Stakeholders of TRUSTEE. Ensuring that any data manipulation and AI functions are deployed on TRUSTEE have respect for fundamental human rights, within a framework underpinned by the rule of law, builds a solid foundation for further introducing and enforcing more abstract ethical principles and values, which guides specifically the deployment of AI. As legally enforceable rights, the fundamental rights fall under the first component of these legal and socio-ethical frameworks and procedures for data manipulation and AI use, making data manipulation and AI use lawful, which safeguards compliance with the law. Drawing from an advanced HRESIA assessment and evaluation of data manipulation and AI use, TRUSTEE's Socio-ethical requirements add an ethical component to the overall framework of requirements (across all the Personas identified in TRUSTEE), therefore making data manipulation and AI use ethical and human-centred. This means that data manipulation and AI deployment on a platform such as TRUSTEE may not be necessarily legally binding, however, the project ensures that such processing is ethical as ethical principles are seen as incremental to ensure trustworthiness of the whole platform of TRUSTEE and its intended AI capabilities.

Several other logics have been guiding the development of the outlined Socio-ethical requirements. The first one reflects the co-design methodology by bringing multi-stakeholder participation to understand the diverse needs and challenges and to consider diversity, non-discrimination, and fairness when designing and implementing AI and data manipulation on open data platforms such as TRUSTEE. The Socio-ethical requirements have adhered to this multi-stakeholder (multi-persona) voice.

⁹⁷⁵ The EU is constitutionally committed to protect the fundamental and indivisible rights of human beings. The Union aims to ensure respect for the rule of law, foster democratic freedom and promote the common good. These rights are articulated in Articles 2 and 3 of the Treaty on European Union, and in the Charter of Fundamental Rights of the EU.

⁹⁷⁶ Council of Europe, Committee of experts on internet intermediaries 2018; European Data Protection Supervisor, Ethics Advisory Group (2018) Towards a digital ethics.
https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf.

A second logic has been to avoid unfair bias by ensuring datasets made available for processing via TRUSTEE does not suffer from inadvertent historic bias, incompleteness, or poor governance. Requirements addressing data providers attend to this logic to prevent the continuity of such biases to lead to unintended perpetuation of prejudice and discrimination against specific groups or individuals. Furthermore, the framework and procedures centre individuals and their fundamental rights, freedoms, and agency. Within the advancing context of data manipulation and AI processing, the freedom of the individual requires mitigation of (in)direct illegitimate coercion, threats to mental autonomy, wellbeing, unjustified surveillance, deception, and unfair manipulation. These conditions are reflected within the Socio-ethical requirements that different TRUSTEE Personas must adhere to. These requirements are also applicable across the entire lifecycle of data manipulation and AI processing and individually address several stakeholders, which will be further defined in the following sections.

A third logic guiding the development of the requirements has been to ensure that there is accessibility and universal design principles guiding the AI functions on TRUSTEE. Specifically, in open data markets where data comes from a wide variety of diverse domains and jurisdictions, and concerning a diverse sector of societies and societal aspects, AI functionalities should be user-centric and human-focused, designed in ways that allow all people to use the functionalities in an intuitive way, regardless of the age, genre, abilities or characteristics of the users of the platforms and without any risks that they can unintentionally inflict harms by deploying AI functions through the Platform. The AI deployed to the TRUSTEE platform should not have a one-size-fit-all approach but consider Universal Design principles⁹⁷⁷ to address the widest possible range of users, as well as following relevant accessibility standards⁹⁷⁸.

Considering that data providers on TRUSTEE are now foreseeing the use of green technologies and sustainable digital futures, their responsibility towards TRUSTEE should fall at the level of sustainability, as a fourth logical requirement. As such, TRUSTEE extends this by supporting the development and use of sustainable digital technologies that have a smaller environmental impact. TRUSTEE seeks to explore data manipulation and AI functionalities that are operationalised through harmonised technical standards.

Transparency obligations also underpin the outlined requirements. The framework positions TRUSTEE to be assumed as high-risk or having certain level of risk in its deployment of AI processing and data manipulation. Therefore, treating it with an assumed level of risk, the framework and procedures guide TRUSTEE Stakeholders to adhere to the principles of transparency and accountability. As such, the development outlines a clear purpose (objective), defined as the reason for data collection and use. For instance, an AI model based on the use of personal data should always be developed, trained, and deployed with a clearly defined purpose. This demonstrates how socio-ethical principles underpin regulatory requirements. While TRUSTEE works with datasets deemed low risk due to the lack of real data and personal data, any repurposing and AI manipulation deployed is seen within a greater context such as future applicability. This is to say that the developed framework and procedures assume possible real data use and real contexts for data manipulation via TRUSTEE.

Lastly, as cybersecurity plays a crucial role in ensuring that the TRUSTEE platform with data manipulation and AI functionalities is resilient against malicious attacks, standardization around data sharing and processing should play a key role to ensure Socio-ethical requirements are met. To this end, the framework procedures look at provisions from both non-technical and technical perspectives and suggest that privacy and security are met at design level and throughout the lifecycle of the project including during the exploitation of data through AI.

OUTCOMES AND ANALYSIS

⁹⁷⁷ See for example EN 301 549

https://www.etsi.org/deliver/etsi_en/301500_301599/301549/03.02.01_60/en_301549v030201p.pdf

⁹⁷⁸ Consider web accessibility standards <https://www.w3.org/WAI/standards-guidelines/wcag/> and the United Nations Convention on the Rights of Persons with Disabilities

Many outcomes were generated using the methods outlined above, which are analysed in the following sub-sections. Personas, User Stories, User-centred Requirements are presented in the first two sub-sections, followed by Legal and Socio-ethical Requirements. Due to the large number of outcomes produced by this methodology, the complete lists of User-centred Requirements, identified Personas and their User Stories, Legal Requirements, and Socio-ethical Requirements are included in Annex 3: Personas and User Stories, Annex 4: User-Centred Requirements and Interconnection with Legal Requirements, Annex 5: Legal Requirements, Annex 6: Socio-Ethical Requirements.

TRUSTEE PERSONAS AND USER STORIES

The distribution of the Pilot Leader and Pilot Member Survey version resulted in the collection of responses from the 6 different data domains corresponding to TRUSTEE Pilot Use cases. Following the gathering and investigation of the responses, three different user groups were formed for each Pilot Use case domain according to the job role of the respondents, which formulated a set of preliminary TRUSTEE Personas for each Pilot Use case, as follows: Manager, Researcher, and General Employee. Each group included several job roles, namely: the Manager user group included job roles such as Project Manager, Technical Manager, etc.; the Researcher user group included job roles such as Associate Researcher, Senior Researcher, etc.; and the General Employee user group included job roles such as Data Scientist/Analyst, Software Developer, AI/ML Engineer/Developer, etc. These preliminary TRUSTEE Personas were then combined and mapped to the different user role groups of the TRUSTEE Platform and produced the finalized TRUSTEE Personas, which relate to the Data Provider User Role, the Data Consumer User Role, and the TRUSTEE Developer User Role. This process is illustrated in Figure 62 below.

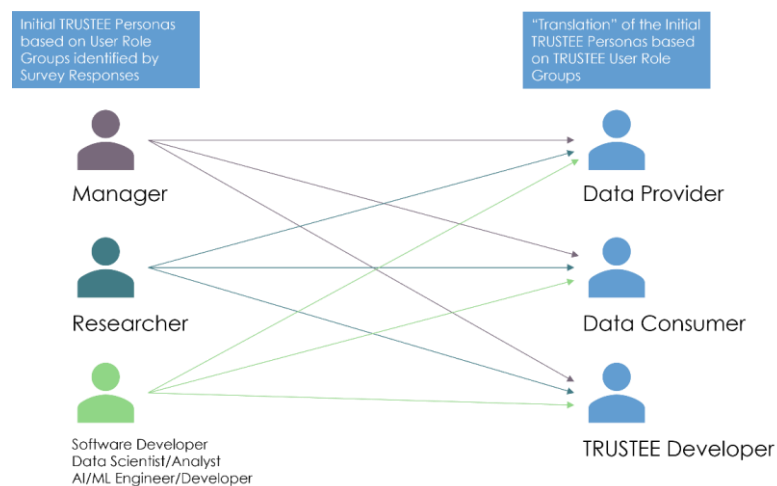


Figure 62: TRUSTEE Survey Personas mapping

The extraction of the goals and frustrations from each Persona produced the User Stories, which, upon evaluation, were gathered and grouped according to the Pilot User case domain and the corresponding TRUSTEE Persona, and eventually led to the extraction of the TRUSTEE User-centred requirements. The distribution of the Pilot Leader and Pilot Member Survey versions resulted in the collection of 35 responses, which created 18 Personas, 3 for each Pilot domain of TRUSTEE. These Personas produced 140 User Stories, which were then analysed and, eventually, formulated 119 User-centred requirements, as it is showcased in Figure 63, below. A complete list of the TRUSTEE Platform Personas, for each domain, and their User Stories is included in Annex 3: Personas and User Stories.

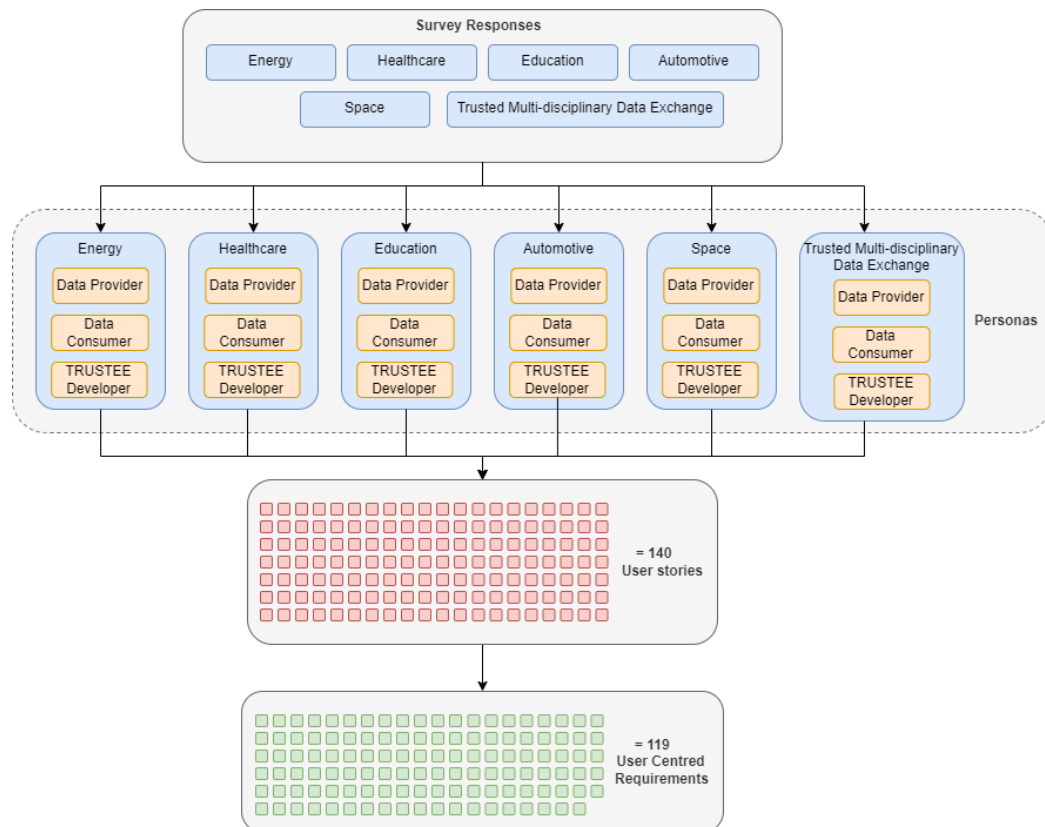


Figure 63: Extracted User-centred requirements from Surveys.

TRUSTEE USER-CENTRED REQUIREMENTS

A User-centred requirement is formulated from a User story, which is extracted from a Persona that was developed based on one or more survey responses. However, it is possible that a single User story produces more than one User-centred requirements while it is also possible that more than one Use stories are combined to produce a single User-centred requirement. The process of extracting User-centred requirements is described in this section by presenting, as an example, the extraction of two User-centred requirements from the Data Provider Persona of the Automotive domain. Figure 64, below, refers to Jim Cross, the TRUSTEE Data Provider Persona created by combining survey responses from the Automotive Pilot Use Case domain.

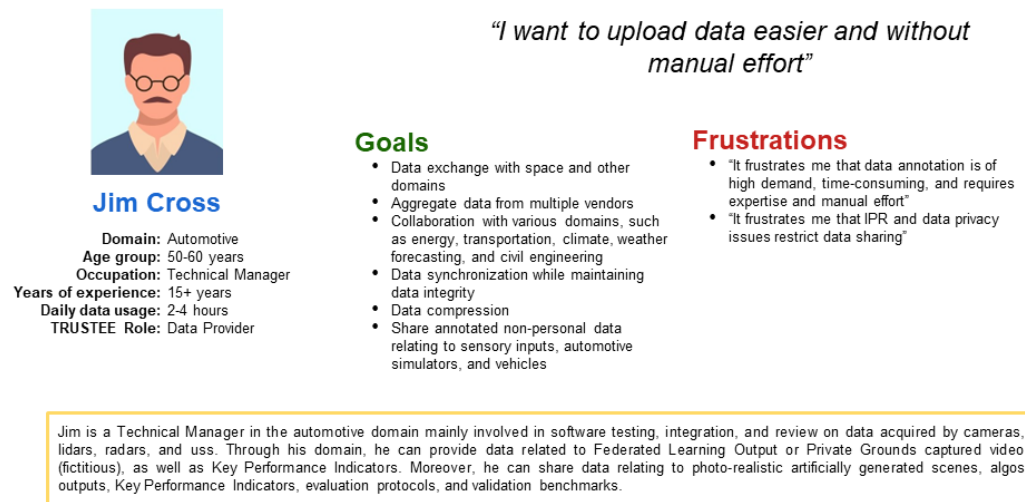


Figure 64: TRUSTEE Data Provider Persona

This Persona produced several User stories that described its goals and frustrations as a Data Provider. One of these User stories describes how Jim wants to be certain about the way his data are used, for which he has provided TRUSTEE with access, so that he can terminate the procedure if he identifies any violations of his interests, as it can be seen in Figure 65.

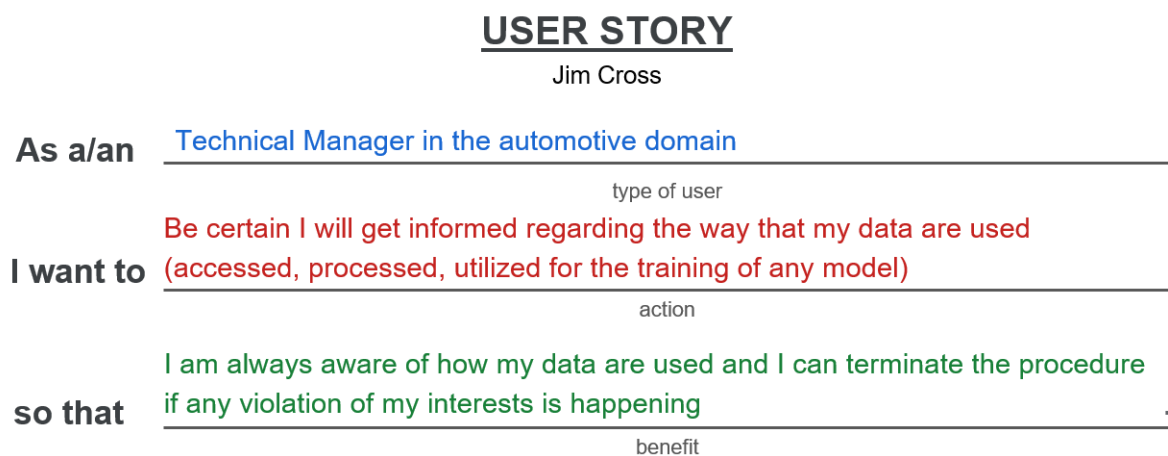


Figure 65: TRUSTEE Data Provider User Story

Eventually, this User story produced 2 User-centred requirements, that were identified as Mandatory and Desirable, respectively, as follows:

- “The user must always be aware of how their data are used and they must be able to terminate the procedure if any violation of their interests may happen”.
- “The user should get informed regarding the way that their data are used (accessed, processed, utilized for the training of any model)”.

This process was followed for all Personas and all User stories for every Pilot Use case of TRUSTEE, in order to extract the User-centred requirements of the TRUSTEE Platform. The priorities of the requirements, namely Mandatory, Desirable, Optional, Future enhancement, were identified based on the phrasing used by the respondents of the TRUSTEE Survey when reporting their needs and objectives. The complete list of the User-centred requirements extracted for the TRUSTEE Platform as well as their interconnection with Legal requirements, which will be explained further below, is included in Annex 4: User-Centred Requirements and Interconnection with Legal Requirements, where for each User-centred Requirement the following information is provided:

- Req ID: Refers to the assigned ID of the User-centred requirement
- Personas: Refers to the Persona/s (Data Provider, Consumer, Model Provider, Developer) in which the requirement relates to
- Requirement: Refers to the description of the requirement, which stems out from the User story/stories
- Priority: Refers to the assigned priority of the User-centred requirement (Mandatory, Desirable, Optional, Possible Future Enhancement)
- Interconnected Legal requirements: Refers to the interconnection of User-centred requirements with the relating Legal requirement/s for TRUSTEE System Developers and for TRUSTEE end-users (personas)
- Architecture Building Block: Refers to the interconnection of User-centred requirements with the ABBs that they relate to, which are described in more detail in Part V: TRUSTEE Architecture.

CROSS-VALIDATION OF USER-CENTRED REQUIREMENTS

The complete list of the TRUSTEE User-centred requirements was validated through Workshop 1.0 and more specifically through the discussions held during the 1st Session.

The 1st Session of the Workshop gathered all participants together, both on-site and online, to discuss on Personas and User-centred requirements and constituted the co-design and cross-validation process regarding User-centred requirements between TRUSTEE partners and external participants coming from the TANGO⁹⁷⁹, TEADAL⁹⁸⁰, AUTOFAIR⁹⁸¹, and GLASS⁹⁸² projects as well as additional external attendees keen on Data Spaces and Data exchange in EU. The 1st Session was divided into two sub-sessions, namely the presentation of Personas and User-centred requirements and the discussion of major findings stemming from the methodologies that were followed, as well as from the User-centred requirements that were extracted. The discussion was complemented by several questions regarding end-user interest in using the TRUSTEE Platform, needs and additional requirements that such a platform should cover and adhere to, foreseen obstacles, categories of data that can be potentially shared, and importance of data characterization and metadata standards. These questions were derived from the TRUSTEE Survey in order to aid the discussion between the participants. The major findings of this discussion, which refer collectively to all Pilot domains of TRUSTEE, and which enhanced the co-design process and led to the cross-validation of the extracted TRUSTEE User-centred requirements, are summarized below:

- All Workshop participants expressed a great interest in using the TRUSTEE Platform as end-users.
- Strong security and identity management are of major importance, mostly for data providers sharing either their data or AI models.

⁹⁷⁹ <https://cordis.europa.eu/project/id/101070052>

⁹⁸⁰ <https://cordis.europa.eu/project/id/101070186>

⁹⁸¹ <https://cordis.europa.eu/project/id/101070568>

⁹⁸² <https://cordis.europa.eu/project/id/959879>

- HE combined with FL was argued as a strong privacy mechanism for sharing AI models and enhancing knowledge extraction through their iterative aggregation.
- Multidisciplinary data exchange and secondary use of data is a primary focus of end-users.
- Privacy restrictions and IPR issues are a major concern of data providers and an obstacle commonly faced by data consumers.
- Data providers want to provide access to their data for fostering cross-domain collaboration; however, without actually sharing their raw data.
- Constant need for alignments between different data formats and difficulties in performing complex data queries hinder data consumers from actually using data for secondary purposes.
- Data quality and validity are major issues for both data providers and data consumers.
- Risk of data loss often prevents data providers from sharing their data.
- Data providers want to keep total control over the data they share.
- Software developers, AI/ML engineers, and Data scientists need a platform where data computations can be performed, use and re-use of AI/ML models can be feasible, and knowledge sharing can be enhanced.
- Proper data characterization may in one hand depend on the Use case (data domain) but is of major importance when considering secondary use of data, while following metadata standards that establish the minimum data characteristics of a dataset should be highly recommended for all actors sharing data in a multi-disciplinary data exchange scenario.

Specifically, regarding the Automotive and the Healthcare domain, additional findings are presented below:

- Automotive: quality and integrity of data are important, as well as real-time requirements, data synchronization for data fusion, and processing of large volumes of data
- Healthcare: strong identity management is important, leading to connecting people with specific intentions, following a formalization of intended use of the data while agreements are always logged somewhere

What is more, access to high volume and diverse data was discussed as a challenge by all Workshop participants due to the difficulty of controlling and ensuring the volume and diversity of the data shared by data providers in a platform such as TRUSTEE. Multi-disciplinary data exchange and secondary use of data, where data are used for a different purpose than their primary purpose of use, were both argued as approaches that can potentially enhance and increase both data volume and diversity. Additionally, collaboration among the projects that participated in the TRUSTEE Workshop, and more specifically synergies between TRUSTEE, TANGO and TEADAL, was discussed and will be explored further as the projects progress, proposing an additional measure towards enhancing secondary data use, thus, increasing data volume and diversity.

Based on the Workshop outcomes and insights gathered, the already identified Personas, User Stories, User-centred requirements for TRUSTEE were not only cross-validated but enhanced as well. An additional Persona, named “Model Provider”, was identified, leading to the extraction of additional User stories and User-centred requirements, as showcased in Figure 66, below. Workshop 1.0 led to the finalization of the TRUSTEE Personas which include: the Data Provider Persona, who wants to register their data and make them accessible through TRUSTEE; the Model Provider Persona, who wants to provide their AI/ML models to TRUSTEE; the Consumer Persona that comprises both Data and Model Consumers; and the Developer Persona, who wants to use several functionalities that are offered through TRUSTEE as part of its incremental strategy.



Figure 66: Finalized TRUSTEE User-centred requirements

TRUSTEE LEGAL REQUIREMENTS

As revealed in the “Methodology” section, the lists of legal requirements has been created separately for (1) the TRUSTEE System Developers, i.e. the entities which are developing TRUSTEE architecture and its building blocks, and will further provide the technical solution to end users, as well as for (2) different types of TRUSTEE end-users, i.e. Data and Model Providers, Consumers (considering both data and model Consumers), and Developers, who would make use of the technology which TRUSTEE will expose as part of its incremental strategy. However, it should be noted that both these lists present a preliminary version of the Legal requirements for the TRUSTEE Platform and will be revisited as the project and technical development progresses.

Below sections describe the manner in which each of the lists should be read by the addressees.

TRUSTEE LEGAL REQUIREMENTS FOR END-USERS

The initial list of Legal requirements for TRUSTEE users is presented in a form of a table, representing the following logic applicable to reading its content. It should be stressed that the structure of the list has been adjusted to the one presented in User-centred requirements:

- Column 1 provides the information about the **requirement identification number (Req ID)**, which reflects the matter of the requirement (legal requirements, “LEG”) and the addressee of the requirement (“USR”, i.e., user of the TRUSTEE solution).
- Column 2 indicates which user exactly is **the addressee of the requirement**: Data Provider, Consumer, Model Provider or Developer, as defined in the Personas and User Stories section above .
- Column 3 provides the **content of the requirement**. The requirements stem from various legal domains, such as: data protection, data privacy, data governance, regulations concerning contractual arrangements, cybersecurity /information security, AI regulations and guidelines, considerations concerning blockchain application, eIDAS Regulation, intellectual property law. All relevant acts, which provide mentioned requirements, are indicated in Part I: State of the Art and the requirements present the initial interpretation of the applicability of obligations and responsibilities included in the legal acts considered in Part I: State of the Art in the TRUSTEE environment.
- Column 4 includes information about **the relevance of the requirement for the user and for TRUSTEE**. Unlike the User-centred requirements, the priority of the Legal requirements do not depend on the users and their perspective but is dependent on the aspect whether a specific legal norm is in force, applicable and binding for the user, and the non-compliance may trigger a liability issue or not. Following this logic, all Legal requirements stemming from the legal acts and frameworks which are adopted and in force have been considered as “Mandatory Requirements”. “Desirable Requirements” are the requirements which TRUSTEE could help the users address (see Open Data Directive, as referenced in the Legal framework and procedures and the list of Legal requirements), as well as the requirements which are extracted from the legal acts being still at the proposal stage or in the law-making process, i.e., not adopted. These requirements may be still subject to further updates, depending on the final shape of the legal act at the moment of its adoption and publication.
- Column 5 provides initial information which **TRUSTEE building blocks should support the user in addressing respective requirements** and in what other manner TRUSTEE could help in responding to it.
- Column 6 highlights the **abbreviation of the legal act or the name of the relevant standard/guideline and the provision** where the requirement has been extracted from.

- Column 7 provides information about the **status of the act from Column 6**, and indicates whether the act is already adopted, in force and applicable, if the applicability is not present yet, or if the act is still in the pre-adoption phase, nevertheless, due to its relevance for the TRUSTEE solution, the requirements included therein have already been indicated and, wherever required, will be further adjusted to the final, form of the act once it is adopted by relevant bodies. The updates will be reflected in D2.2 and D2.3.
- Column 8 – this column includes **additional comments**, which may help the users and TRUSTEE System Developers to navigate which requirements are pre-existing for the users even before their attempt to use TRUSTEE and should be addressed by them with or without TRUSTEE support.

In Annex 4: User-Centred Requirements and Interconnection with Legal Requirements, the User-centred requirements have been combined with Legal requirements, to showcase the interplay of both types of requirements - in other words, to show that in the process of fulfilment of a specific User-centred requirement, additional Legal requirements have to be taken into account and addressed by the TRUSTEE architecture, to enable data exchange and implementation of the processes and data flow compliant with laws applicable to the TRUSTEE environment, from the perspective of the TRUSTEE System Developers as well as TRUSTEE end-users. Currently, only “Mandatory” and “Desirable” User-centred requirements have been considered in the aforementioned mapping between Legal and User-centred requirements. Requirements of these priorities will be the initial focus during the development of the TRUSTEE Platform. This mapping may be expanded to include Optional User-centred requirements as the project progresses, if necessary. These modifications will be included in future versions of this deliverable, i.e., D2.2 and D2.3.

TRUSTEE LEGAL REQUIREMENTS FOR TRUSTEE SYSTEM DEVELOPERS

The initial list of Legal requirements for TRUSTEE System Developers is presented in a form of a table, representing the following logic applicable to reading its content. It should be stressed that the structure of the list has been adjusted to the one presented in User-centred requirements:

- Column 1 provides the information about the **requirement identification number (Req ID)**, which reflects the matter of the requirement (Legal requirements, “LEG”) and the addressee of the requirement (“TSD”, i.e., TRUSTEE System Developer, meaning TRUSTEE Partners developing and further maintaining TRUSTEE solution), as well as the number of the requirement.
- Column 2 provides the **content of the requirement**. The requirements stem from various legal domains, such as: data protection, data privacy, data governance, regulations concerning contractual arrangements, cybersecurity /information security, AI regulations and guidelines, considerations concerning blockchain application, eIDAS Regulation, intellectual property law. All relevant acts, which provide mentioned requirements, are indicated in Part I: State of the Art, and the requirements present the initial interpretation of the applicability of obligations and responsibilities included in the legal acts considered in Part I: State of the Art in the TRUSTEE environment.
- Column 3 includes the information about **the relevance of the requirement for the user and for TRUSTEE**. Unlike the User-centred requirements, the priority of the Legal requirements do not depend on the users and their perspective but is dependent on the aspect whether a specific legal norm is in force, applicable and binding for the user, and the non-compliance may trigger a liability issue or not. Following this logic, all Legal requirements stemming from the legal acts and frameworks which are adopted and in force have been considered as “Mandatory Requirements”. “Desirable Requirements” are the requirements:
 - which TRUSTEE System Developers should consider, as these would allow them to fulfil the mandatory requirements in a more appropriate or expected manner,
 - which are extracted from the legal acts being still at the proposal stage or in the law-making process, i.e., not adopted. These requirements may be still subject to further updates, depending on the final shape of the legal act at the moment of its adoption and publication.

- Requirements stemming from relevant guidelines or pieces of literature, which are not legal acts per se and hence, are not mandatory by default, nevertheless, should be considered by the TRUSTEE System Developers.
- Column 4 provides initial information which **TRUSTEE building blocks should support the user in addressing respective requirement** and in what other manner TRUSTEE could help in responding to it.
- Column 5 highlights the **abbreviation of the legal act or the name of the relevant standard/guideline and the provision** where the requirement has been extracted from.
- Column 6 provides the information about the **status of the legal act /standard or relevant guideline from Column 5**, and indicates whether the act is already adopted, in force and applicable, if the applicability is not present yet, or if the act is still in the pre-adoption phase, nevertheless, due to its relevance for the TRUSTEE solution, the requirements included therein have already been indicated and, wherever required, will be further adjusted to the final, form of the act once it is adopted by relevant bodies. The updates will be reflected in D2.2. and D2.3.

Following the example of linking User-centred requirements, considered as “Mandatory” and “Desirable”, with the Legal requirements addressed to the TRUSTEE users, Annex 4: User-Centred Requirements and Interconnection with Legal Requirements also includes a similar linking between “Mandatory” and “Desirable” User-centred requirements and Legal requirements for TRUSTEE system developers.

TRUSTEE SOCIO-ETHICAL REQUIREMENTS

Similar to the approach followed for developing the initial list of the Legal Requirements for TRUSTEE, two lists for the Socio-ethical requirements have been developed, namely: for the TRUSTEE System Developers, i.e. the entities which are developing TRUSTEE architecture and its building blocks, and will further provide the technical solution to end users; the different types of TRUSTEE end-users, i.e. Data and Model Providers (providing data or AI/ML models to TRUSTEE), Consumers (considering both data and model Consumers), and Developers (who would make use of the technology which TRUSTEE will expose as part of its incremental strategy). However, it should also be mentioned for the Socio-ethical requirements as well, that both these lists present a preliminary version of the Socio-ethical requirements for the TRUSTEE Platform and will be revisited as the project and technical development progresses. Updates and enhancements to these lists will be provided in upcoming version of this deliverable, namely D2.2 and D2.3.

PRIVACY AND SOCIO-ETHICAL TOOLKIT

Train, Validate, Test

When compiling a training dataset containing personal data, some precautions must be taken to ensure compliance with the regulations. Checks should be carried out on the origin of the data, on the actual possibility of reusing it for training purposes or on the measures taken to limit the risks of misuse.

The quality of the algorithm's output is closely linked to the quality of the training dataset, regardless of the categories of the concerned data. Some criteria such as data accuracy, distribution need to be checked to limit the

risk of error when using the algorithm, especially when it has consequences for individuals. A checklist is provided by CNIL for such checking⁹⁸³.

PRIVACY BY DESIGN

Privacy by Design is related to concerns to be predominant throughout the whole design, development, and implementation process of software components. In the context of AI systems, it involves providing explicit consent from data owners before using their data in training the model and respecting clauses like the “right to be forgotten” under General Data Protection Regulation (GDPR), whenever invoked by the data owner. First solutions for implementing such clauses in AI systems has started to be proposed and largely remains unexplored^{984 985 986}.

FEDERATED LEARNING ARCHITECTURE

FL architectures enable training AI systems, mainly neural network models, without collecting users’ personal data at a centralised location. With FL, users do not have to share data, rather only training model updates. Differential Privacy⁹⁸⁷ and secure aggregation⁹⁸⁸ enforce another layer of privacy to make training updates more secure. When using the FL models, the TRUSTEE System Developers must check that the data used in local is independent and identically distributed (a condition ensuring that the information drawn from the data will reflect the same trends without being specific to each centre)⁹⁸⁹.

EXAMINE BIASES

The bias is well known problem in the ML domain for data scientists. It is an anomaly in the output of ML models, due to the prejudiced assumptions made during the model development process or prejudices in the training data. The risks of discrimination linked to the use of an algorithm trained on biased data are widely recognized. However,

⁹⁸³ <https://www.cnil.fr/en/collecting-and-qualifying-training-data>

⁹⁸⁴ Yinzhi Cao, Junfeng Yang: Towards Making Systems Forget with Machine Unlearning. IEEE Symposium on Security and Privacy 2015: 463-480

⁹⁸⁵ D. Barua, “A time to remember, a time to forget: user controlled, scalable, life long user modelling,” The University of Sydney, Australia, 2016.

⁹⁸⁶ Abhishek Kumar, Tristan Braud, Sasu Tarkoma, Pan Hui: Trustworthy AI in the Age of Pervasive Computing and Big Data. PerCom Workshops 2020: 1-6

⁹⁸⁷ C. Dwork, A. Roth et al., “The algorithmic foundations of differential privacy,” Foundations and Trends® in Theoretical Computer Science, vol. 9, no. 3–4, pp. 211–407, 2014.

⁹⁸⁸ K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, “Practical secure aggregation for privacy-preserving machine learning,” in Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 2017.

⁹⁸⁹ <https://www.cnil.fr/en/collecting-and-qualifying-training-data>

the factors contributing to the risk of biases remain and the methods to correct them are in their infancy. The training dataset must therefore be thoroughly inspected for signs of potential bias⁹⁹⁰.

TRUSTEE SOCIO-ETHICAL REQUIREMENTS FOR TRUSTEE SYSTEM DEVELOPERS

The initial list of socio-ethical requirements for TRUSTEE users is presented in a form of a table, representing the following logic applicable to reading its content. It should be stressed that the structure of the list has been adjusted to the one presented in Annex 5: Legal Requirements.

The Socio-ethical requirements underpin the main legal laws and principles that have guided this work (further detail in D1.6⁹⁹¹ and Annex 6: Socio-Ethical Requirements). Table 6.1 lists the socio-ethical requirements for TRUSTEE System Developers. It should be read as follows: The first column addresses the nature of requirement; the second presents the requirement ID; the third provides a requirement description; the fourth assigns a priority to the requirement based on the legal act or relevant guideline/standard from which the requirement is stemming; the fifth refers to the legal act or relevant guideline/standard; and the last column offers a mapping between the Socio-ethical requirement and the Legal requirements for TRUSTEE, which sends the reader to Annex 5: Legal Requirements.

The Socio-ethical requirements were broadly split into six categories:

- **Fundamental human rights and freedoms:** Adhering to the fundamental human rights and respect for human dignity, autonomy, and freedom. That is, any data manipulation and algorithmic models deployed on TRUSTEE should ensure freedom to the individual (human freedom to make life decisions for themselves), by providing clear choices and alternatives to the data manipulation and AI functions offered on TRUSTEE, and other requirements within this arch.
- **Privacy-by-design, security-by-design, lawful-by-design for transparency, trust, and accountability:** requirements include ensuring universal designs for accessibility, explicability, and explainability; ensuring evaluation and assessment phases of AI and data manipulation, ensuring clear security standards, data privacy mechanisms, and others.
- **Prevent harm:** requirements include preventing harm through technological security and safety of the systems that handle data and data processing; ensuring that there are proper measures to assess risks and safety; providing mitigation plan for any foreseeable risks of data manipulation.
- **Ethical requirements through technical and non-technical implementation:** addressing TRUSTEE System Developers, these requirements look at both technical and non-technical implementation of ethical principles, including providing opportunities for stakeholder participation and eliciting regular feedback; providing and updating codes of conduct, but also ensuring that the system itself is clear about the regulatory obligations with regards to data manipulation. As such, the socio-ethical principles systematically make direct reference to legal guardrails and rules and ensures that ethical principles are not observed separately from Legal requirements.

⁹⁹⁰ <https://www.cnil.fr/en/collecting-and-qualifying-training-data>

⁹⁹¹ Spanakis, E., Sakkalis, V.-F., Forgó, N., Barud, K., Özkayagan-Prändl, H., Sailllant, C. – U., Markakis, E., Papatsaroucha, D. – H., Rosengren, P. – C., Kourtis, A., Xilouris, G. – Z., Hillman, V. – E., Nicolin, A. I. – I., Niskanen, I. – V., Martignano, M.-U., RUTTO, G., TESFAI OGBU, I.-R., Grguric, A. – E., & Lalos, A. – I. (2022). *TRUSTEE Consortium (2022) D1.6 Data management plan available*.

- **Communication, documentation, and reporting for accountability transparency:** Achieving trustworthiness requires continuous and systematic communication, documentation, reporting, and updating of protocols around the system and its functionalities. The requirement within this arch seek to ensure that whether developer, data or model provider ensure that there are clear communication channels around the characteristics, limitations and potential shortcomings of the platform, the data available on it, and the AI functionalities.
- **Assessment of ethical designs and practices:** any ethical principles require reiterative process of re-evaluation and updates as technological innovation continues to accelerate. Keys stakeholders of TRUSTEE must provide regular assessment and Socio-ethical requirements throughout the lifecycle of the project; ensure that prior to ethical and trustworthy assessments the TRUSTEE System Developers also meet the legally binding requirements, whenever and whichever relevant, and that trustworthiness is not an end in itself but rather an ongoing process of assessment, clear responsibilities, and good governance.

Full list of the Socio-ethical requirements for TRUSTEE System Developers is included in Annex 6: Socio-Ethical Requirements: 6.1 Socio-ethical Requirements for TRUSTEE System Developers.

TRUSTEE SOCIO-ETHICAL REQUIREMENTS FOR TRUSTEE END-USERS

The initial list of Socio-ethical requirements for TRUSTEE users is presented in a form of a table, representing the following logic applicable to reading its content. It should be stressed that the structure of the list has been adjusted to the one presented in Annex 5: Legal Requirements - 5.1 LEGAL REQUIREMENTS FOR TRUSTEE END-USERS. The table should be read as follows: first column addresses the nature of the requirement; the second presents the requirement ID; the third mentions the Persona (Data Provider, Model Provider, Consumer, Developer) that the requirement refers to; the fourth provides a description of the requirement; the fifth presents the priority of the requirement based on the legal act or relevant guideline standard from which the requirement is stemming; the sixth indicates the legal act relevant guideline/standard mapped to the requirement; and the last column offers a mapping between the Socio-ethical requirement and the Legal requirements for TRUSTEE, which sends the reader to Annex 5: Legal Requirements .

On a similar vein to the Socio-ethical requirements addressing TRUSTEE System Developers, the Socio-ethical requirements addressing TRUSTEE end-users adhere to (initially) two overarching sets. This is not to say that users of TRUSTEE should not meet or even be aware of any of the other criteria. The focus in this first iteration is on the foundational principles of meeting requirements that aim to protect the fundamental human rights and freedoms of individuals and the collective and adhere to the requirements around documentation, transparency, accountability, and reporting in relation to data processing and the use of TRUSTEE's AI functionalities. Specifically:

- **Fundamental human rights and freedoms:** Consumers (both data and model consumers) should clearly be informed about, know, and respect the fundamental human rights of those whose data is shared on TRUSTEE and be fully aware, know and respect the conditions around how such data use and processing may be impacting individuals and the collective. Consumers should be clearly informed about the platform's or its AI models' and data limitations and how TRUSTEE meets these same fundamental rights.
- **Communication, documentation, and reporting for accountability transparency:** Equally, a range of ethical requirements address end-users of TRUSTEE's AI functionalities and available data by seeking that they provide transparency around the AI and data manipulation as functions with capabilities as well as limitations which don't act as replacement of or on human individuals' behalf; that they ensure societal and environmental wellbeing through explicability, and transparency around choice and alternative ways of achieving the same outcomes; that they document and communicate benefits and opportunities from trustworthy open data platforms such as TRUSTEE but also their limitations.

Full list of the requirements is in Annex 6: Socio-Ethical Requirements.

PART IV: TECHNICAL REQUIREMENTS

PART IV SUMMARY

This section is dedicated to the systematic elicitation and definition of the functional and non-functional requirements for the TRUSTEE project, in alignment with the objectives set out in WP2. Leveraging the 5-step Agile Approach⁹⁹², the focus is on systematically identifying the necessary requirements for the areas of data management, processing and analytics, security and trust across distributed and federated data sets, privacy, legal and ethical issues, homomorphic encryption, self-sovereign identity management, explainable and trustworthy AI, and FAIR and open knowledge extraction.

The methodology adopted during this phase of eliciting and defining requirements was a dynamic and iterative one, as necessitated by the Agile Approach. It began with a vision statement defining the overall aim of the TRUSTEE project, followed by brainstorming sessions to generate potential requirements. Extensive surveys as well as a workshop with stakeholders were conducted to capture user needs and expectations, as presented in **Error! Reference source not found..** This was further reinforced by tapping into the partners of the TRUSTEE consortium's collective expertise and exploring best practices from their network of collaborators.

After generating potential requirements, these were organized and prioritized based on their value and alignment with the project's goals. These requirements were then formalized and regularly reviewed and revised in line with Agile's principle of iterative refinement.

This deliverable represents a living document and will be updated regularly throughout the life of WP2 in a truly Agile manner. As the project evolves, some requirements may change, new ones may emerge, and others may become obsolete. Therefore, updates, refinements, and extensions will be made as necessary to ensure the requirements remain relevant and aligned with the project's goals.

The results of this requirement elicitation and definition phase serve as a blueprint for the development of innovative solutions across the entire TRUSTEE ecosystem. They provide a clear, prioritized roadmap for the team to follow, ensuring that the solutions developed are useful, usable, and valuable to the end-users and meet the project's objectives of creating a secure, trustworthy, and privacy-aware data federation framework.

PART IV INTRODUCTION

The challenge of eliciting and defining technical requirements for complex digital systems is well-recognized in the industry. Key considerations during this process include ensuring comprehensive coverage of system functionalities, aligning the solution with stakeholder expectations, and planning for contingencies. Industry standards and best practices advocate for systematic, structured approaches to requirement elicitation, allowing for iterative refinement and adaptation to changes.

For the TRUSTEE project, which aims to contribute significantly to the European Data Space's technological architecture, the choice of methodology for requirement elicitation was of paramount importance. After a careful evaluation of various industry standards and methodologies, TRUSTEE opted for the 5-step Agile Approach to define the technical functional and non-functional requirements of its Application Building Blocks. These Building Blocks include Homomorphic Enabled Data Fusion (HEDF); AI Models as a Service (AIMaaS); Accountable Transactions Recorder (ATR); Trustworthy AI Assessment Support Design Framework (TAI-SDF); Knowledge

⁹⁹² Schmitt, Anna, and Sabrina Hörner. "Systematic literature review–improving business processes by implementing agile." *Business Process Management Journal* 27, no. 3 (2021): 868-882.

Repository (KR); Authentication Manager (AM); Data Protection Impact Assessment (DPIA); One Stop Shop (OneSS); SSI Homomorphic Enabled Framework (SSI-HE); Data Lake (DL); Security & Trust Manager (STM); and TRUSTEE Dashboard (DA), which will be further described in **Error! Reference source not found.**

The 5-step Agile Approach was chosen due to its adaptability, iterative nature, and its emphasis on collaboration and continuous improvement - values that align well with TRUSTEE's objectives and the dynamic landscape of the digital data market. It allows the project to remain flexible and adaptable, swiftly incorporating changes and feedback, and continuously refining the requirements as more information becomes available or conditions change. The remainder of this deliverable is structured as follows:

- TRUSTEE Application Building Blocks Technical Requirements
 - Methodology for Collecting Technical Requirements
 - 5-Step Agile Approach
 - Adopting 5-Step Agile Approach in TRUSTEE
 - List of User-centred requirements mapped into ABBs and translated into Technical requirements

The sections that follow will delve into the specifics of how the Agile Approach was adopted in TRUSTEE and the resulting Technical requirements for each of the TRUSTEE ABBs. This approach ensures the development of innovative solutions that are in line with the needs of the stakeholders and the overall objectives of the European Data Space.

TRUSTEE APPLICATION BUILDING BLOCKS TECHICAL REQUIREMENTS

The TRUSTEE vision on the secure and trustworthy data sharing ecosystem, proposes an all-encompassing framework that follows the lifecycle of secure-by-design cloud services from the development to the deployment and operation stages. To make sure that the analysis phase will capture all appropriate requirements for this broad architecture, the project has put special effort to define and follow a structured technical requirement gathering methodology, directly mapped and sprout from User-centred requirements already gathered from extended focused surveys and a stakeholder's workshop.

METHODOLOGY FOR COLLECTING TECHNICAL REQUIREMENTS

The requirements of TRUSTEE form the bedrock upon which this complex project is established. They provide a definitive direction and vision that will permeate all aspects of the project, from architectural choices to financial directives. Crafting concise, clear requirements is a challenging task; if mishandled, it could lead to serious consequences. For instance, the project could significantly overshoot its budget, or the resulting product might not meet expectations, fail to serve a useful purpose, or not function as intended - circumstances that are all undesirable. More distressingly, it could breach the promise of being secure, trustworthy, and privacy-aware - the cornerstones of the TRUSTEE initiative. Therefore, due to the project's complexity and multidimensional nature, a systematic and adaptable approach, the "5-step Agile Approach" has been selected as the methodology to collect the technical requirements of the visioned TRUSTEE solutions.

5-STEP AGILE APPROACH

The 5-step Agile Approach is a known industry standard methodology for requirements solicitation that provides a user-centric, collaborative, flexible, risk-reducing, continually improving, and quality-assured framework that supports the complex needs of technical projects such as TRUSTEE. It enables the team to navigate the challenges of creating a secure, trustworthy, and privacy-aware data federation framework in a rapidly evolving technological landscape. For the TRUSTEE development team, the methodology is particularly beneficial for a variety of reasons:

1. **User-Centric:** The Agile Approach is fundamentally user-centric, which means it's designed to prioritize user needs. TRUSTEE has already conducted extensive survey campaigns and stakeholder workshops; thus, the team has a strong understanding of the users' needs, desires, and pain points. This user-centric approach ensures that the developed solutions are useful, usable, and valuable to the end-users.
2. **Collaborative:** The Agile Approach encourages collaboration. TRUSTEE development team, pilot owners, and stakeholders continually interact throughout the development process. This approach prevents silos, promotes the sharing of ideas, and helps to align everyone to the project goals.
3. **Flexible and Responsive:** The Agile methodology is iterative, which allows the team to be flexible and responsive to changes. Throughout TRUSTEE lifetime stakeholder needs change, new technologies emerge, or the team finds a better way to meet a requirement; the Agile Approach allows them to adjust the project scope and direction accordingly. This adaptability is crucial in the ever-evolving field of cloud and blockchain technologies, which TRUSTEE aims to develop.
4. **Risk Mitigation:** Agile reduces the risk of complete project failure. By delivering the project in increments, the team can ensure that the most important features are developed first. Even if later stages of the project face delays or issues, the team would have already delivered value to the stakeholders.
5. **Continuous Improvement:** Agile encourages a culture of continuous improvement. After each iteration, the team reflects on what went well and what could be improved. These retrospectives help the team to continually refine their process, making them more efficient and effective over time.
6. **Quality Assurance:** Agile emphasizes sustainable development, which includes regular and thorough testing throughout the project. This built-in quality assurance helps the team to identify and fix issues quickly and helps to ensure the reliability and security of the developed solutions – a crucial aspect for TRUSTEE, given its focus on privacy and security.

ADOPTING A 5-STEP AGILE APPROACH IN TRUSTEE

Through the 5-step Agile Approach, the TRUSTEE project can effectively gather, prioritize, and refine the technical and non-technical requirements necessary to build a secure, trustworthy, and privacy-aware data federation framework. This methodology, combined with extensive user surveys and stakeholder workshops, ensures that the project's solutions meet user needs and fulfil the vision of the project. TRUSTEE's 5-step agile approach is analysed as follows:

1. **Define TRUSTEE Vision:** The TRUSTEE project already has a clear vision: a secure, trustworthy, and privacy-aware data federation framework. Utilizing the 5-step Agile Approach, this vision is communicated to the team and guides the extraction and creation of the user story backlog. This step ensures that all subsequent requirements gathering, and prioritization efforts align with the primary goals of the project.
2. **Brainstorm:** In the context of TRUSTEE, brainstorming occurred following user surveys and stakeholder workshops. The ideas, pain points, and features brought up during these sessions have been used as a springboard for additional brainstorming within the project team. This step helps identify potential functional requirements (like the need for homomorphic encryption or self-sovereign technologies using Blockchain) and non-functional requirements (such as system performance, security, scalability, and availability).
3. **Organize and Prioritize:** After generating a list of potential Technical requirements, these were organized and prioritized. Again, the user surveys and stakeholder workshops acted as significant source of input for this step. User-centred requirements identified as critical by stakeholders have been given high priority. This step is crucial to ensure that the most valuable features are developed first and that the work aligns with stakeholders' needs and expectations.
4. **Formalize:** In the formalize step, the team has documented the final list of prioritized Technical requirements, creating a more detailed specification for each one. Each requirement has enough detail to allow the development team to understand what needs to be built and the QA team to understand how to

test it. At the same time, Agile encourages avoiding over-specification early on, leaving room for the team to find the best solution during development.

5. **Review and Revise:** The iterative nature of Agile means that TRUSTEE development team will continually review and revise the Technical requirements based on stakeholder feedback, new discoveries during development, and changes in the project environment (such as new technologies or changes in the data federation space). This step helps the project to stay flexible and adapt to changing needs or conditions.

Through biweekly calls with all consortium partners beginning early in the project, TRUSTEE also made sure that the "Right People" were convened and that "Everyone in the Room" was engaged. The project's cloud repository and a collaboratively edited excel file in a type-checked organised manner have been used to aid brainstorming in the "Record Information" process. The required description was analytically developed as shown in Table 0.14, and champions for each technical issue undertook the 'Review and Revise' procedure to guarantee the accuracy and completeness for each case. The goal of this task was not to produce an exhaustive and disorderly list of technical details, but rather to highlight the main requirements that would jolt the design phase into coming up with the best solutions to fulfil them.

Table 0.14 Requirements Description Template

	Template Field	Description
S	ID	Req-TASK-x (unique id) Where: - TASK = Domain (DA AM HEDF SSI-HE ATR DPIA TAI-SDF AIMaaS HEDF KR DL OneSS) - x = FUNC NFUNCIIdentifier.ShortDesc Example: "Req-AM-FUNC-1"
	Dependencies	Requirement ID of the related User-centred requirement(s) that this Technical requirement is based on.
	Type	Choose from the list: - FUNC: Functional Suitability - DATA: Data - USE: Usability - REL: Reliability - SEC: Security - PERF: Performance Efficiency - COMP: Compatibility - MAINT: Maintainability - PORT: Portability
	Short name	Meaningful and not too long
	Personas	The TRUSTEE Personas involved in this Technical requirement
	Description	General description
	Additional Information	Any additional information pertinent to the specific Technical requirement

	Priority (MoSCoW)	This allows to identify the priority of the requirements; it can be updated in the different iterations: • M: Must-have. Mandatory requirement • S: Should-have. Desirable requirement • C: Could-have. Optional requirement • W: Will-not-have. Possible future enhancement
M	Measures	Strong encouragement for defining concrete measures. Means of measuring the goal achievement. Milestones, metrics, etc.
A	Achievable	How achievable is the requirement? This will have to consider the Objectives, the Priority, and any possible foreseen obstacles.
R	Objectives	Objectives of this requirement. Why is the goal worthwhile? What needs does it fulfil and how relevant are they to the project? Does it satisfy any future dependencies (to be filled with the requirement IDs)?
T	Timeline	Strong encouragement for defining concrete timelines (Such as milestones, delivery deadlines according to work plan, etc.). When should this requirement be achieved? This goes together with the future dependencies and their timelines.

The following section lists the functional and non-functional requirements identified by applying the above methodology for each TRUSTEE ABB,

- AI Models as a Service (AIMaaS)
- Accountable Transactions Recorder (ATR)
- Authentication Manager (AM)
- Data Protection Impact Assessment (DPIA)
- Homomorphic Enabled Data Fusion (HEDF)
- Data Lake (DL)
- One Stop Shop (OneSS)
- SSI Homomorphic Enabled Framework (SSI-HE)
- Security & Trust Manager (STM)
- Knowledge Repository (KR)
- Trustworthy AI Support Design Framework (TAI-SDF)
- TRUSTEE Dashboard (DA)

AI MODELS AS A SERVICE (AIMAAS)

	Template Field	Description
S	ID	Req-AIMaaS-FUNC-1
	Dependencies	USR-031, USR-052, USR-062, USR-085, USR-095
	Type	FUNC: Functional Suitability

	Short name	xAI incorporation
	Personas	Consumer, Model Provider
	Description	The AIMaaS tool should include explainable AI functions.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The AIMaaS tool should explain how the AI process generates the additional knowledge extracted from data.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AIMaaS-FUNC-2
	Dependencies	USR-052, USR-081, USR-085, USR-114, USR-115
	Type	FUNC: Functional Suitability
	Short name	Synthetic data generation
	Personas	Developer, Consumer
	Description	The AIMaaS tool could provide a model for the generation of synthetic data, based on real-world data.
	Additional Information	None
	Priority (MoSCoW)	C: Could-have. Optional requirement
M	Measures	Implementation and validation

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TRUSTEE users could be able to generate artificial data based on real-world diversity.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AIMaaS-FUNC-3
	Dependencies	USR-052, USR-089, USR-093, USR-115
	Type	FUNC: Functional Suitability
	Short name	Fast and easy model deployment
	Personas	Developer, Consumer, Model Provider
	Description	The AIMaaS tool should utilize a containerized approach to provide the deployment and run of AI/ML models in an easy and fast way.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TRUSTEE user should be able to extract knowledge without complex analysis methods and tools
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AIMaaS-FUNC-4
	Dependencies	USR-053, USR-062, USR-085
	Type	FUNC: Functional Suitability
	Short name	Reuse of trained models
	Personas	Model Provider, Consumer
	Description	The TRUSTEE platform should store trained AI/ML models for the users to reuse.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE users should be able to reuse already trained AI/ML models/algorithms.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AIMaaS-FUNC-5
	Dependencies	USR-052, USR-085, USR-089, USR-092
	Type	PERF: Performance Efficiency
	Short name	Improved AI methods
	Personas	Developer, Model Provider, Consumer
	Description	The AIMaaS tool could enable the Model Provider to share self-developed improved AI/ML models for the Consumer and Developer to use.

	Additional Information	None
	Priority (MoSCoW)	C: Could-have. Optional requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The AIMaaS tool could support improved AI/ML analysis methods and tools.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AIMaaS-FUNC-6
	Dependencies	USR-093
	Type	REL: Reliability
	Short name	Direct connection with TAI-SDF
	Personas	Developer, Model Provider, Consumer
	Description	The AIMaaS tool could redirect the users of AI models to the TAI-SDF, ensuring the use of trustworthy AI models.
	Additional Information	None
	Priority (MoSCoW)	C: Could-have. Optional requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The AIMaaS tool could support fully trustworthy AI/ML models.
T	Timeline	Year 1: independent test of functionality

		Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AIMaaS-NFUNC-1
	Dependencies	USR-052, USR-062, USR-085
	Type	USE: usability
	Short name	User-friendly AIMaaS tool
	Personas	Developer, Consumer, Model Provider
	Description	The AIMaaS tool should have an easy-to-use and efficient UI.
	Additional Information	AIMaaS is planned to have its own user interface to facilitate specific requests regarding pre-trained AI models availability, or AI/ML training results regarding dataset shared by users.
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The AIMaaS should offer a user-friendly way to interact with it and extract knowledge.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AIMaaS-NFUNC-2
	Dependencies	USR-052, USR-062, USR-085, USR-093
	Type	USE: usability

	Short name	Results presentation
	Personas	Developer, Consumer
	Description	The AIMaaS tool should present results in a user-friendly way.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The AIMaaS tool should include good and user-friendly tools, in order for the users to present efficiently the data analysis results.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

ACCOUNTABLE TRANSACTIONS RECORDER (ATR)

	Template Field	Description
S	ID	Req-ATR-FUNC-1
	Dependencies	USR-004, USR-005, USR-045
	Type	FUNC: Functional Suitability
	Short name	Decentralised network of Nodes
	Personas	Data Provider, Model Provider, Consumer
	Description	The decentralised network will consist of nodes each with distinct functionalities. The peers play a crucial role in validating, endorsing, and committing transactions. The nodes also take part in ledger management.
	Additional Information	ATR building block is an integral part of the integrity and accountability framework promised by TRUSTEE and as such it has dependencies with User-centred requirements relevant to secure and privacy preserving data sharing.

	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Every user/Persona from outside the blockchain network must be connected to a peer node to read/write from the ledger.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-ATR-FUNC-2
	Dependencies	USR-057, USR-005, USR-045
	Type	FUNC: Functional Suitability
	Short name	Ordering Service
	Personas	Data Provider, Model Provider, Consumer
	Description	Ordering service creates a consistent order of transactions before transactions are placed in the blockchain. The ordering service also ensures that all nodes comprise the same version of the ledger.
	Additional Information	ATR building block is an integral part of the integrity and accountability framework promised by TRUSTEE and as such it has dependencies with User-centred requirements relevant to secure and privacy preserving data sharing.
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The ordering nodes are responsible for the consensus process.

T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-ATR-FUNC-3
	Dependencies	USR-004, USR-005, USR-039, USR-045, USR-057
	Type	FUNC: Functional Suitability
	Short name	Membership Service Provider (MSP)
	Personas	Data Provider, Model Provider, Consumer
	Description	FUNC: Functional Suitability
	Additional Information	A component should be in place which is responsible for managing all the digital identities of the blockchain network which includes every user/Persona that will be connected to ATR.
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	A Membership Service Provider (MSP) ensures that only authorised participants can access and participate in the network. MSP defines policies for managing identity registration, authentication, authorization and so on.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-ATR-FUNC-4

	Dependencies	USR-004, USR-005, USR-045, USR-057
	Type	FUNC: Functional Suitability
	Short name	Chaincode
	Actors	Data Provider, Model Provider, Consumer
	Description	The chaincode is responsible for executing all the smart contracts in the network. It contains the business logic and defines how the infrastructure is going to behave, and what types of transactions will take place.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Chaincode defines the rules and conditions that must be met in order for transactions to be validated. It can be designed to ensure privacy, for instance, private data collection, and it is executed in a secured environment to prevent malicious actors from tampering with the code.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-ATR-FUNC-5
	Dependencies	USR-001, USR-004, USR-005, USR-006, UAR-022, USR-039, USR-043, USR-057
	Type	USE: Usability
	Short name	Blockchain Monitoring Interface
	Personas	Data Provider, Consumer, Developer
	Description	Hyperledger Explorer, an open-source tool will be integrated to the blockchain.

	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Hyperledger Explorer provides an interface to explore and interact with the blockchain. It is a web interface which provides graphical representation of the blockchain network. It can be utilised to determine transaction details, monitor performance, track network activity, etc.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-ATR-NFUNC-1
	Dependencies	UAR-005, USR-006, USR-039, UAR-045, USR-057, USR-064
	Type	SEC: Security, PERF: Performance Efficiency
	Short name	Decentralised network Benchmarking
	Personas	Data Provider, Model Provider, Consumer
	Description	Hyperledger Caliper is a benchmarking tool to be integrated with the decentralised network to measure the performance.
	Additional Information	This is a security performance requirement
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.

R	Objectives	The main objective is to monitor the performance to detect anomalies. The integration with Hyperledger caliper will enable measuring all successful and failed transactions, transaction and read latency, transaction and read throughput, resource consumption, etc.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

AUTHENTICATION MANAGER (AM)

	Template Field	Description
S	ID	Req-AM-FUNC-1
	Dependencies	USR-003, USR-005, USR-008, USR-039, USR-57, USR-100
	Type	FUNC: Functional Suitability
	Short name	Self-Sovereign User Identity
	Personas	Data Provider, Model Provider, Consumer
	Description	TRUSTEE users/Personas have control of their identity and the data they provide to the platform.
	Additional Information	Although AM is not specifically depended on “USR-040: TRUSTEE should provide protection of personal data”, the SSI can help in protecting personal data and minimise the information a user has to share.
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE users/Personas have control of their identity and the data they provide to the platform, to protect their rights based on EU regulations, and increase the sense of security and transparency.

T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AM-FUNC-2
	Dependencies	USR-005, USR-008, USR-011, USR-019, USR-039, USR-044, USR-054, USR-057, USR-100
	Type	FUNC: Functional Suitability
	Short name	Secure User Authentication
	Personas	Data Provider, Model Provider, Consumer
	Description	To access the various systems users/Personas will have to be authenticated.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	To access the various systems users will have to be authenticated, in order to increase security.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AM-FUNC-3
	Dependencies	USR-005, USR-008, USR-011, USR-039
	Type	FUNC: Functional Suitability

	Short name	Secure User Registration
	Personas	Data Provider, Model Provider, Consumer
	Description	The system will have to allow the registration of new users/Personas.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The system will have to allow the registration of new users, in order for them to be authorized to access the system after getting authenticated.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AM-FUNC-4
	Dependencies	USR-005, USR-039
	Type	SEC: Security
	Short name	Strong Authentication of Users
	Personas	Data Provider, Model Provider, Consumer
	Description	TRUSTEE should provide strong user authentication mechanics.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement.
M	Measures	Implementation and validation.

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE should provide strong user authentication mechanics to ensure cybersecurity.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AM-FUNC-5
	Dependencies	USR-003, USR-019, USR-044
	Type	SEC: Security
	Short name	Trusted User Provided Data
	Personas	Consumer, Data Provider
	Description	TRUSTEE should ensure the information provided by Data Providers and Consumers is trusted.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE should ensure the information provided by the user/Persona is trusted.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AM-NFUNC-1
	Dependencies	USR-003, USR-019, USR-054, USR-057
	Type	PERF: Performance Efficiency
	Short name	Efficient User Authentication
	Personas	Data Provider, Model Provider, Consumer
	Description	TRUSTEE should provide efficient user authentication mechanics.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE should feature an efficient authentication mechanism to increase performance and enhance the user experience.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-AM-NFUNC-2
	Dependencies	USR-003, USR-019, USR-054, USR-057
	Type	USE: Usability
	Short name	Efficient Authentication of Users/Personas
	Personas	Data Provider, Model Provider, Consumer
	Description	TRUSTEE should feature an easy-to-use authentication.
	Additional Information	None

	Priority (MoSCoW)	S: Should-have. Desirable requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE should feature an easy-to-use authentication to facilitate authentication and enhance the user experience.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

DATA PROTECTION IMPACT ASSESSMENT (DPIA)

	Template Field	Description
S	ID	Req-DPIA-FUNC-1
	Dependencies	USR-001, USR-003, USR-005, USR-006, USR-029, USR-033, USR-046, USR-057
	Type	SEC: Security
	Short name	Assessment based on CNIL methodology
	Personas	Consumer, Data Provider, Developer
	Description	The tool MUST allow the user to carry out the risk assessment based on CNIL methodology.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The tool MUST allow the user to carry out the risk assessment based on CNIL methodology, to enhance security.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DPIA-FUNC-2
	Dependencies	USR-001, USR-007, USR-005, USR-009, USR-010, USR-029, USR-036, USR-038, USR-040, USR-041, USR-042, USR-045, USR-054, USR-100,
	Type	SEC: Security
	Short name	Cross component data communication assessment
	Personas	Consumer, Data Provider, Developer
	Description	The DPIA tool MUST assess the privacy impact of the data communication(s) between the TRUSTEE components.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The DPIA tool MUST assess the privacy impact of the data communication(s) between the TRUSTEE components, in order to protect data privacy.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DPIA-FUNC-3
	Dependencies	USR-005, USR-042
	Type	SEC: Security
	Short name	Abide to EU legal knowledge base
	Personas	Data Provider, Consumer
	Description	The DPIA tool MUST be provided from an EU member public authority and abide to the legal knowledge base of the EU.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The DPIA/PIA tool MUST be provided from an EU member public authority and abide to the legal knowledge base of the EU, in order to ensure proper distribution and compliance with the EU directives.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DPIA-FUNC-4
	Dependencies	USR-006, USR-007, USR-010, USR-040, USR-045, USR-107
	Type	SEC: Security
	Short name	Assess purpose of processing data
	Personas	Data Provider, Consumer

	Description	The tool MUST include the module allowing to describe the purpose of processing the data, and, where applicable to the user, legitimate interest in processing.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The tool MUST include the module allowing to describe the purpose of processing the data, and, where applicable to the user, legitimate interest in processing, to enhance user privacy and give the user control of their data.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

HOMOMORPHIC ENABLED DATA FUSION (HEDF)

	Template Field	Description
S	ID	Req-HEDF-FUNC-1
	Dependencies	USR-001, USR-002, USR-003, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-035, USR-048, USR-077, USR-087, USR-095, USR-110, USR-111, USR-114
	Type	DATA: Data
	Short name	Data import from multiple sources /domains
	Personas	Consumer, Data provider
	Description	TRUSTEE should be able to accept data from various sources and domains and fuse them based on common elements.
	Additional Information	
	Priority (MoSCoW)	M: Must-have. Mandatory requirement

M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The HEDF ABB should be able to accept and process data from various sources/domains and fuse them based on their common attributes, according to their semantic representations.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNCT-2
	Dependencies	USR-001, USR-003, USR-005, USR-009, USR-029, USR-033, USR-036, USR-038, USR-041, USR-042, USR-045, USR-046, USR-057, USR-107
	Type	DATA: Data
	Short name	Data privacy
	Personas	Data provider
	Description	Data provided to TRUSTEE should always be homomorphically encrypted. Data will be processed in encrypted form.
	Additional Information	
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.

R	Objectives	To ensure security, privacy and compliance with relevant laws (GDPR etc.) and ethical obligations, data within the TRUSTEE ecosystem will always be in encrypted form. The data itself will never be decrypted. The result of the processing which will contain no personally identifiable information will be decrypted and securely sent to the user.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNCT-3
	Dependencies	USR-001, USR-002, USR-003, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-035, USR-048, USR-077, USR-087, USR-095, USR-110, USR-111, USR-114
	Type	DATA: Data
	Short name	Data import in various formats
	Personas	Data provider
	Description	TRUSTEE could be able to accept data in different formats without requiring the data provider to reformat them.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Every data field (e.g., a date) can be represented in different formats across the various data sources. TRUSTEE could be able to accept most formats without requiring that the data provider adheres to a strict format guideline. It will then be able to process and fuse data regardless of the format used.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNCT-4
	Dependencies	USR-001, USR-002, USR-003, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-035, USR-048, USR-077, USR-087, USR-095, USR-110, USR-111, USR-114
	Type	DATA: Data
	Short name	Compatibility with existing multi-disciplinary data domains
	Personas	Data provider
	Description	TRUSTEE should be able to catalogue and retrieve data from existing infrastructures that host data from various domains.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	In addition to being able to retrieve data from domain specific data sources, TRUSTEE could be able to interface with infrastructures that hold or catalogue multi-disciplinary data. In this case, these infrastructures may act as intermediaries between TRUSTEE and the actual data provider.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNCT-5
	Dependencies	USR-001, USR-002, USR-003, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-035, USR-048, USR-077, USR-087, USR-095, USR-110, USR-111, USR-114
	Type	FUNC: Functional Suitability

	Short name	Data computation and analysis
	Personas	Consumer
	Description	TRUSTEE should be able to perform computations and analysis over the data selected by the user.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE should be able to not only fuse data from different sources but also perform user selected computations on them. In addition, data transformations or reformats could also be needed.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DL-FUNCT-6
	Dependencies	USR-001, USR-002, USR-003, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-035, USR-048, USR-077, USR-087, USR-095, USR-110, USR-111, USR-114
	Type	PERF: Performance Efficiency
	Short name	Efficient data querying
	Personas	Consumer
	Description	The TRUSTEE platform should be able to search across multiple data sources quickly and efficiently for the user selected criteria.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE will catalogue a very large volume of data from different sources and in different formats. Searching through all that data should be done with efficiency and speed to facilitate the best use of the project infrastructure's resources as well as a desirable user experience.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNC-7
	Dependencies	USR-001, USR-002, USR-003, USR-005, USR-009, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-033, USR-035, USR-036, USR-038, USR-041, USR-042, USR-045, USR-046, USR-047, USR-048, USR-050, USR-052, USR-057, USR-064, USR-067, USR-068, USR-075, USR-077, USR-081, USR-087, USR-088, USR-094, USR-095, USR-096, USR-098, USR-101, USR-102, USR-104, USR-107, USR-110, USR-111, USR-112, USR-113, USR-114, USR-115, USR-116
	Type	FUNC: Functional reliability
	Short name	Basic FL
	Personas	Data Provider, Consumer, Model Provider, Developer
	Description	FL categories that must be supported: fedAvg algorithm with a simple weighted average fusion rule, with synchronous operation among participating nodes.
	Additional Information	-
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.

R	Objectives	The TRUSTEE HEDF will be able to support the basic federated learning paradigm to TRUSTEE users
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNC-8
	Dependencies	USR-001, USR-002, USR-003, USR-005, USR-009, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-033, USR-035, USR-036, USR-038, USR-041, USR-042, USR-045, USR-046, USR-047, USR-048, USR-050, USR-052, USR-057, USR-064, USR-067, USR-068, USR-075, USR-077, USR-081, USR-087, USR-088, USR-094, USR-095, USR-096, USR-098, USR-101, USR-102, USR-104, USR-107, USR-110, USR-111, USR-112, USR-113, USR-114, USR-115, USR-116
	Type	FUNC: Functional reliability
	Short name	Non-iid FL
	Personas	Data Provider, Consumer, Model Provider, Developer
	Description	FL categories that should be supported: Cluster-based FL
	Additional Information	-
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TRUSTEE HEDF will be able to support the federated learning paradigm over non-iid datasets to TRUSTEE users
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNC-9
	Dependencies	USR-003, USR-029, USR-047, USR-064, USR-067, USR-068, USR-075, USR-077, USR-081, USR-094
	Type	FUNC: Functional reliability
	Short name	Broadcasting global NN for FL
	Personas	Data Provider, Consumer, Model Provider, Developer
	Description	HE-enabled FL framework should be able to broadcast global NN model in an uncoded fashion to participating clients
	Additional Information	-
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	This requirement enables the actual implementation of the FL framework via TRUSTEE
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNC-10
	Dependencies	USR-001, USR-002, USR-003, USR-005, USR-009, USR-012, USR-013, USR-016, USR-018, USR-029, USR-032, USR-033, USR-035, USR-036, USR-038, USR-041, USR-042, USR-045, USR-046, USR-047, USR-048, USR-050, USR-052, USR-057, USR-064, USR-067, USR-068, USR-075, USR-077, USR-081, USR-087, USR-088, USR-094, USR-095, USR-096, USR-098, USR-101, USR-102, USR-104, USR-107, USR-110, USR-111, USR-112, USR-113, USR-114, USR-115, USR-116
	Type	FUNC: Functional reliability

	Short name	Entire NN FL
	Personas	Data Provider, Consumer, Model Provider, Developer
	Description	HE-enabled FL framework should support the update of entire NN models among participating agents.
	Additional Information	-
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	This requirement enables the actual implementation of the basic FL framework
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-HEDF-FUNC-11
	Dependencies	USR-001, USR-002, USR-003, USR-005, USR-009, USR-012, USR-013, USR-016, USR-018, USR-029,USR-032,USR-033,USR-035,USR-036,USR-038,USR-041,USR-042,USR-045,USR-046,USR-047,USR-048,USR-050,USR-052,USR-057,USR-064,USR-067,USR-068,USR-075,USR-077,USR-081,USR-087,USR-088,USR-094,USR-095,USR-096,USR-098,USR-101,USR-102,USR-104,USR-107,USR-110,USR-111,USR-112,USR-113,USR-114,USR-115,USR-116
	Type	FUNC: Functional reliability
	Short name	Personalized FL
	Personas	Data Provider, Consumer, Model Provider, Developer
	Description	HE-enabled FL framework should support the update of targeted NN model parts (e.g., only specific layers) among participating agents.
	Additional Information	-

	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	This requirement enables the actual implementation of the basic FL framework
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

DATA LAKE (DL)

	Template Field	Description
S	ID	Req-DL-FUNC-1
	Dependencies	USR-001, USR-002, USR-003, USR-012, USR-013, USR-016, USR-017, USR-18, USR-20, USR-021, USR-022, USR-023, USR-025, USR-026, USR-027, USR-028, USR-029, USR-032, USR-034, USR-035, USR-036, USR-048, USR-058, USR-060, USR-069, USR-070, USR-071, USR-073, USR-074, USR-075, USR-078, USR-079, USR-081, USR-083, USR-086, USR-087, USR-088, USR-090, USR-094, USR-099, USR-102, USR-103, USR-104, USR-106, USR-112, USR-113, USR-114, USR-115
	Type	FUNC: Functional Suitability
	Short name	Efficient data querying
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE platform should be designed to perform quick and efficient searches across multiple data sources based on the user's selected criteria.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement

M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE will catalogue a very large volume of data from different sources and in different formats. Searching through all that data should be done with efficiency and speed to facilitate the best use of the project infrastructure's resources as well as a desirable user experience.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DL-FUNC-2
	Dependencies	USR-001, USR-003, USR-029, USR-36, USR-112
	Type	DATA: Data
	Short name	Data processing in accordance with privacy regulations
	Personas	Data Provider, Model Provider, Consumer
	Description	All data processing within TRUSTEE should be done according to relevant privacy regulations. No personally identifiable data should ever be revealed.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	TRUSTEE will index data that contains personally identifiable information. Data should be handled in such a way that this information is never revealed to any user or even to the TRUSTEE platform itself. TRUSTEE Consumers will receive processed data that do not contain any such information.

T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DL-FUNC-3
	Dependencies	USR-001, USR-003, USR-019, USR-020, USR-029, USR-36, USR-103, USR-112
	Type	DATA: Data
	Short name	Data processing in accordance with data security and IPR regulations
	Personas	Data Provider, Model Provider, Consumer
	Description	TRUSTEE should comply with relevant IPR and data security regulations. Data providers should be able to control how their data is processed and for which users.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Data within TRUSTEE should always be transferred and processed in a secure way. Data providers should be able to control which users and process their data and which types of processing they are allowed to perform. Specific licensing agreements should be used for any data provider to endure proper data usage.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

ONE STOP SHOP (ONESS)

	Template Field	Description
S	ID	Req-OneSS-NFUNC-StdAPIs
	Dependencies	USR-014, USR-017, USR-063, USR-066
	Type	COMP: Compatibility
	Short name	APIStandard
	Personas	Developer
	Description	APIs and connection points of TRUSTEE ABBs shall be compliant with existing standards.
	Additional Information	APIs based on Standards
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	External interface specification inspection from D2.1, implementation inspection between ABBs owners (endpoint providers and endpoint consumers).
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Integration effectiveness and efficiency standardized and industry practices, usual way of working. Adoption of best practices
T	Timeline	D2.1 part V: Architecture ABB external interfaces in M12, M12 delivery of initial version of implementation, milestones to corresponds with implementation delivery and pilot milestones. .

	Template Field	Description
S	ID	Req-OneSS-NFUNC-StdAPIDoc
	Dependencies	USR-014, USR-017, USR-51, USR-058, USR-061, USR-063, USR-066, USR-108
	Type	COMP: Compatibility
	Short name	APIStandard documentation
	Personas	Developer

	Description	APIs and connection points of TRUSTEE ABBs shall be documented according to existing established standards (e.g. OpenAPI for RESTful interfaces) using available open source tools.
	Additional Information	APIs documentation based on Standards
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	External interface specification inspection from D2.1, implementation inspection between Building Blocks owners (endpoint providers and endpoint consumers).
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Integration effectiveness and efficiency standardized and industry practices, usual way of working. Adoption of best practices
T	Timeline	D2.1 part V: Architecture ABB external interfaces in M12, M12 delivery of initial version of implementation, milestones to corresponds with implementation delivery and pilot milestones.

SSI HOMOMORPHIC ENABLED FRAMEWORK (SSI-HE)

	Template Field	Description
S	ID	Req-SSI-HE-FUNC-1
	Dependencies	USR-001, USR-002, USR-013, USR-020, USR-026, USR-027, USR-048, USR-064
	Type	FUNC: Functional Suitability
	Short name	Data characterization tool
	Personas	Data Provider, Model Provider, Consumer
	Description	Provide a secure tool for data sets characterization/metadata definition.
	Additional Information	This tool includes the characteristics definition (credential issuance), check (credential storage and visualization) and verification (credential validation).
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.

M	Measures	Specific tool implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	This objective provides a way to data providers to prove their data-sets features in a secure and trustworthy way to data consumers. This functionality is very useful prior to the data fusion in order to securely identify what data to use for a specific purpose (based on the datasets features).
T	Timeline	Year 1: isolated data characterization solution Year 2: integrated data characterization solution in TRUSTEE Year 3: validated data characterization solution

	Template Field	Description
S	ID	Req-SSI-HE-FUNC-2
	Dependencies	USR-001, USR-002, USR-016, USR-020, USR-026, USR-027, USR-032, USR-036, USR-048, USR-050, USR-064, USR-086
	Type	FUNC: Functional Suitability
	Short name	Data characterization schemas
	Personas	Data provider, Model Provider
	Description	The data sets characterization/metadata definition tool must allow the definition of different data sets characterization schemas.
	Additional Information	As many schemas as needed could be created.
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Specific tool implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Each data set (use case) requires different metadata/features to be proven. Different metadata schemas will be defined based on the specific use-case requirements.

T	Timeline	Year 1: isolated characterization schemas definition solution Year 2: integrated characterization schemas definition solution in the data characterization solution (see Req-SSI-HE-FUNC-1). Year 3: validated characterization schemas definition solution

	Template Field	Description
S	ID	Req-SSI-HE-FUNC-3
	Dependencies	USR-016, USR-020, USR-041, USR-042, USR-046, USR-057, USR-104, USR-107
	Type	FUNC: Functional Suitability
	Short name	Confidential metadata
	Personas	Data Provider, Model Provider, Consumer
	Description	Extend the data characterization tool functionality allowing private metadata/characteristics to be validated in a confidential way.
	Additional Information	HE encryption will be used for the metadata/characteristic's confidentiality.
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Specific tool implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Some of the characteristics/metadata features could be private. Encrypted metadata also protects the datasets features.
T	Timeline	Year 1: isolated HE encryption/decryption solution. Year 2: integrated HE encryption/decryption solution in the data characterization solution (see Req-SSI-HE-FUNC-1). Year 3: validated HE encryption/decryption solution.

	Template Field	Description
S	ID	Req-SSI-HE-FUNC-4

	Dependencies	USR-016, USR-020, USR-041, USR-042, USR-046, USR-057, USR-104, USR-107
	Type	FUNC: Functional Suitability
	Short name	Confidential searches over metadata
	Personas	Consumer
	Description	Allow encrypted searches over confidential datasets metadata/attributes for specific queries.
	Additional Information	Searches over encrypted dataset attributes will be implemented on the HE domain.
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Specific tool implementation and validation
A	Achievable	KPI: number of domains (from DoA)
R	Objectives	This objective extends Req-SSI-HE-FUNC-4 by providing a way to verify desired confidential metadata/characteristics of the datasets to be used in TRUSTEE. This functionality is very useful prior to the data fusion in order to privately and securely identify what data to use for a specific purpose (based on the confidential -encrypted- datasets features).
T	Timeline	Year 1: isolated HE searches functionality. Year 2: integrated HE searches functionality in the data characterization solution (see Req-SSI-HE-FUNC-1 and Req-SSI-HE-FUNC-3). Year 3: validated HE searches functionality.

	Template Field	Description
S	ID	Req-SSI-HE-NFUNC-1
	Dependencies	USR-001, USR-002, USR-026, USR-084
	Type	DATA: Data
	Short name	String data characterization
	Personas	Data Provider, Model Provider, Consumer
	Description	All type of metadata/attributes should be supported.
	Additional Information	Numeric values and strings
	Priority (MoSCoW)	S: Should-have. Desirable requirement.

M	Measures	Specific tool implementation and validation KPI: number of domains (different kind of data) (from DoA)
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide HE searchable functionality over numeric values as well as strings/text.
T	Timeline	This requirement will be considered from the beginning of the confidential data characterization solution.

	Template Field	Description
S	ID	Req-SSI-HE-NFUNC-2
	Dependencies	USR-050
	Type	USE: Usability
	Short name	User-friendly GUI for metadata
	Personas	Data Provider, Model Provider, Consumer
	Description	Data characterization solution should include web-based GUI easy to understand by non-technical users.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement.
M	Measures	Specific tool implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a user-friendly web-based GUI for dealing with data characterization issues (for issuance, verification and searching functionalities).
T	Timeline	Year 1: not so relevant. Year 2: preliminary version Year 3: validated final version.

	Template Field	Description
S	ID	Req-STM-FUNC-1
	Dependencies	USR-001, USR-007, USR-029, USR-054, USR-057, USR-064, USR-110
	Type	FUNC: Functional Suitability
	Short name	User-friendly GUI for facilitating collaborations across multi-disciplinary.
	Personas	Data Provider, Consumer, Model Provider
	Description	It offers collaboration across multi-disciplinary data domains for multiple domains: Energy, Health, Education, Automotive, Space, Multi-Disciplinary.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a user-friendly web-based GUI to facilitate the collaborations across multi-disciplinary.
T	Timeline	Year 1: preliminary version with one pilot use case. Year 2: validated version with all domains. Year 3: validated, integrated with the rest of components and tested final version.

	Template Field	Description
S	ID	Req-STM-FUNC-2
	Dependencies	USR-001, USR-007, USR-020, USR-029, USR-036, USR-054, USR-057, USR-064, USR-084, USR-094, USR-110
	Type	FUNC: Functional Suitability

	Short name	Efficient data request
	Personas	Data Provider, Consumer, Model Provider
	Description	It plays a role in data request.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a user-friendly web-based GUI to facilitate the search and request data.
T	Timeline	Year 1: preliminary version for a standalone testing. Year 2: Integration and testing with HEDF. Year 3: validated the final version.

	Template Field	Description
S	ID	Req-STM-FUNC-3
	Dependencies	USR-003, USR-006, USR-020, USR-033, USR-036, USR-043, USR-064, USR-082
	Type	FUNC: Functional Suitability
	Short name	Inform the user about the use of its data
	Actors	Data Provider, Consumers, Model Providers
	Description	It plays a role in the efficiency of data querying.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a user-friendly web-based GUI to inform the data (AI model) provider to write the agreement on how their data (AI model) should be used.
T	Timeline	Year 1: data provider agreement. Year 2: AI model agreement. Year 3: final version.

	Template Field	Description
S	ID	Req-STM-FUNC-4
	Dependencies	USR-001, USR-003, USR-005, USR-006, USR-007, USR-009, USR-010, USR-019, USR-029, USR-033, USR-036, USR-038, USR-040, USR-041, USR-042, USR-045, USR-046, USR-054, USR-057, USR-100, USR-107
	Type	FUNC: Functional Suitability
	Short name	Permissions based on the agreement.
	Personas	Consumer, Data Provider, Model Provider
	Description	It provides a permission to data access based on the agreements and DPIA.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a backend checker tool for checking whether a user may or may not use the requested data and AI model based on the agreement and DPIA.

T	Timeline	Year 1: not provided in the first year. Waiting a preliminary version of DPIA and mainly the DPIA outputs. Year 2: checker permission development and validation. Year 3: integration and final version.

	Template Field	Description
S	ID	Req-STM-FUNC-5
	Dependencies	USR-003, USR-006, USR-020, USR-033, USR-036, USR-043, USR-064, USR-082
	Type	FUNC: Functional Suitability
	Short name	AI model agreement
	Personas	Data Provider, Consumers, Model Providers
	Description	Create an AI model agreement.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a backend side such as a NoSQL database with REST API for creating an agreement for an AI model.
T	Timeline	Year 1: not provided in the first year. Year 2: validated and tested version. Year 3: integration and final version.

	Template Field	Description
S	ID	Req-STM-FUNC-6
	Dependencies	USR-003, USR-006, USR-020, USR-033, USR-036, USR-043, USR-064, USR-082

	Type	FUNC: Functional Suitability
	Short name	Data agreement
	Personas	Consumer, Data Provider, Model Provider
	Description	Create a data agreement
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a user-friendly web-based GUI and a backend side such as a NoSQL database with REST API for creating data agreement.
T	Timeline	Year 1: in the first year. Year 2: validated and tested version. Year 3: integration and final version.

	Template Field	Description
S	ID	Req-STM-FUNC-7
	Dependencies	USR-003, USR-006, USR-020, USR-033, USR-036, USR-043, USR-064, USR-082
	Type	FUNC: Functional Suitability
	Short name	Agreement checker for data fusion
	Personas	Consumer, Data Provider, Model Provider
	Description	Checking the compatibility of the agreements from multiple data providers before the data fusion.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide an agreement checker tool.
T	Timeline	Year 1: not in the first year. Year 2: first and tested version. Year 3: integration and final version.

	Template Field	Description
S	ID	Req-STM-FUNC-8
	Dependencies	USR-003, USR-005, USR-006, USR-020, USR-033, USR-036, USR-043, USR-064, USR-082
	Type	FUNC: Functional Suitability
	Short name	Agreement monitoring
	Personas	Data Provider, Consumer, Model provider
	Description	Monitoring the agreements
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a monitoring tool for showing the agreements, their status (signed/accepted), updated.
T	Timeline	Year 1: not in the first year. Year 2: first and tested version. Year 3: integration and final version.

	Template Field	Description
S	ID	Req-STM-FUNC-9
	Dependencies	USR-001, USR-003, USR-005, USR-006, USR-007, USR-009. USR-010, USR-019, USR-029, USR-033, USR-036, USR-038, USR-040, USR-041, USR-042, USR-045, USR-046, USR-054, USR-057, USR-100, USR-107
	Type	FUNC: Functional Suitability
	Short name	GDPR requirements
	Personas	Data Provider, Consumer, Model provider
	Description	The agreements are created based on the GDPR requirements.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Follow the GDPR requirements from End-to-End of agreement life cycle.
T	Timeline	Year 1: preliminary version in the first year. Year 2: tested version on the use cases. Year 3: integration and final version.

	Template Field	Description
S	ID	Req-STM-NFUNC-1
	Dependencies	USR-001, USR-003, USR-005, USR-006, USR-007, USR-009. USR-010, USR-019, USR-029, USR-033, USR-036, USR-038, USR-040, USR-041, USR-042, USR-045, USR-046, USR-054, USR-057, USR-100, USR-107
	Type	USE: Usability
	Short name	User-friendly GUI for metadata
	Personas	Data provider, Model Provider, Consumer

	Description	The meta-data is used with respect to privacy regulations, legal frameworks and ethical considerations.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	The validation with the legal will be defined.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide a user-friendly web-based GUI for dealing with data characteristics. The metadata will be provided by the data provider and used for searching the data by the data consumer for the requests.
T	Timeline	Year 1: preliminary version. Year 2: validated and integrated version Year 3: final version.

	Template Field	Description
S	ID	Req-STM-NFUNC-2
	Dependencies	USR-001, USR-007, USR-029, USR-054, USR-057, USR-064, USR-110
	Type	USE: Usability
	Short name	User-friendly GUI for facilitating the data request
	Personas	Data provider, Model Provider, Consumer
	Description	It facilitates to search and request a data in a user-friendly way.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation, validation, and testing.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.

R	Objectives	Provide a user-friendly web-based GUI to facilitate the data request.
T	Timeline	Year 1: preliminary version with one pilot use case. Year 2: validated version with all domains. Year 3: validated, integrated with the rest of components and tested final version.

KNOWLEDGE REPOSITORY (KR)

	Template Field	Description
S	ID	Req-KR-FUNC-1
	Dependencies	USR-015, USR-020, USR-022, USR-028, USR-029, USR-030, USR-058, USR-118
	Type	FUNC: Functional Suitability
	Short name	Decentralised network of Nodes
	Personas	Consumer, Data Provider, Model Provider
	Description	The Knowledge repository will be based on IPFS network which will consist of several decentralised nodes. The nodes will be the building blocks of the network and play a crucial role as they will be responsible for storing all datasets in the peer-to-peer network.
	Additional Information	
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Every user/Persona must be connected to the IPFS network to retrieve the stored information from peer nodes.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-KR-FUNC-2
	Dependencies	USR-015, USR-020, USR-029, USR-030, USR-036, USR-052, USR-058, USR-111, USR-118
	Type	USE: Usability
	Short name	Monitoring & Analysis Interface
	Personas	Consumer, Data Provider, Model Provider
	Description	IPFS exporter is a tool which will be integrated with IPFS to export data from an IPFS node to various format, such as JSON or CSV.
	Additional Information	
	Priority (MoSCoW)	M: Must-have. Mandatory requirement.
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The IPFS exporter will enable monitoring and analysing IPFS network activity.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-KR-FUNC-3
	Dependencies	USR-015, USR-020, USR-029, USR-030, USR-036, USR-052, USR-058, USR-111, USR-118
	Type	USE: Usability
	Short name	Data Orchestration
	Personas	Consumer, Data Provider, Model Provider

	Description	IPFS cluster provides data orchestration across a swarm of IPFS daemons.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation.
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	IPFS cluster is a distributed application which will be integrated with the Knowledge Repository (IPFS Network). It will work as a sidecar to IPFS peers, maintaining a global cluster pin set and allocating its items to IPFS peers.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-KR-FUNC-4
	Dependencies	USR-015, USR-020, USR-022, USR-028, USR-029, USR-030, USR-058, USR-118
	Type	USE: Usability
	Short name	Automated Analysis
	Personas	Consumer, Data Provider, Model Provider
	Description	IPFS network does not provide automated analysis functionalities. Thus, other tools and platforms should be integrated to utilise the automated analysis functionality for the stored datasets in IPFS.
	Additional Information	
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation.

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Data stored in the knowledge repo (IPFS network) might be accessed by Apache Spark, etc.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

TRUSTWORTHY AI SUPPORT DESIGN FRAMEWORK (TAI-SDF)

	Template Field	Description
S	ID	Req-TAI-SDF-FUNC-1
	Dependencies	USR-037, USR-091, USR-093
	Type	FUNC: Functional Suitability
	Short name	Phases and principles matrix
	Personas	Consumer, Model Provider, Developer
	Description	The TAI-SDF MUST allow the relative TRUSTEE users/personas to define a matrix for phases and principles, assign, modify, delete such phases and principles following the methodology that the framework will implement, in order to ensure the trustworthiness of the AI models.
	Additional Information	-
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and Validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TAI-SDF will have configuration capabilities

T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-TAI-SDF-FUNC-2
	Dependencies	USR-037, USR-091, USR-093
	Type	FUNC: Functional Suitability
	Short name	Templates and Questionnaires creation
	Personas	Consumer, Model Provider, Developer
	Description	The TAI-SDF MUST support the creation, modification, deletion of relevant templates and questionnaires regarding the validation of trustworthiness that of the AI models and associate them with the phase and principles matrix.
	Additional Information	-
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and Validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TAI-SDF will have the capability to define and configure AI model trustworthiness templates and questionnaires to be associated with the phases and principles matrices.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-TAI-SDF-FUNC-3
	Dependencies	USR-037, USR-091, USR-093

	Type	FUNC: Functional Suitability
	Short name	TAI-SDF project
	Personas	Consumer, Model Provider, Developer
	Description	The TAI-SDF MUST communicate to the TRUSTEE users/personas that a project related to the AI model under development has been created and allow access to information such as: the creation date, last change saved, which is the phase and principle currently under evaluation in the project, so that the TRUSTEE user/persona can select to start or continue an assessment and visualise the assessment output and the received score at each stage along with recommendations provided by the tool for improving the score.
	Additional Information	-
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and Validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TAI-SDF will have the capability to support the user in the assessment
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-TAI-SDF-NFUNC-1
	Dependencies	USR-005, USR-043, USR-057
	Type	Port: Portability
	Short name	User Interface
	Personas	Consumer, Model Provider, Developer
	Description	The TAI-SDF SHOULD be a web application.
	Additional Information	-
	Priority (MoSCoW)	M: Must-have. Mandatory requirement

M	Measures	Implementation and Validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TAI-SDF application should be securely accessible via the web over common web browsers.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

TRUSTEE DASHBOARD (DA)

	Template Field	Description
S	ID	Req-DA-FUNC-1
	Dependencies	USR-049, USR-050, USR-051
	Type	USE: Usability
	Short name	Filtering functionality
	Personas	Consumer
	Description	The TRUSTEE Platform COULD offer a filter functionality for AI/ML models/algorithms and for data domain selection.
	Additional Information	None
	Priority (MoSCoW)	C: Could-have. Optional requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TRUSTEE Dashboard to provide a filtered search to the user for AI/ML models and data domains.

T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-2
	Dependencies	USR-047, USR-49, USR-054, USR-055, USR-59, USR-062, USR-063
	Type	USE: Usability
	Short name	Results Visualization
	Personas	Consumer
	Description	The TRUSTEE Platform SHOULD offer visualization of the results stemming from the users' requests (e.g., search queries, etc.).
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The results from users' requests to be visualized by the dashboard.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-3
	Dependencies	USR-047, USR-49, USR-054, USR-056, USR-59, USR-062, USR-063
	Type	USE: Usability

	Short name	Display of user's activity history.
	Personas	Data Provider. Model Provider. Consumer
	Description	The TRUSTEE Platform COULD provide the user with the ability to request the history of their activity.
	Additional Information	None
	Priority (MoSCoW)	C: Could-have. Optional requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The user to be able to view their history of activity.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-4
	Dependencies	USR-014, USR-047, USR-49, USR-052, USR-054, USR-056, USR-59, USR-060, USR-061, USR-062, USR-063
	Type	USE: Usability
	Short name	Visualize ABBs APIs
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Dashboard MUST visualize the APIs of the building blocks that need to expose information to the Dashboard (e.g., ATR, HEDF for data computation results, etc.).
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and validation

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The APIs from various ABBs to be displayed onto Dashboard.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-5
	Dependencies	USR-059, USR-061
	Type	FUNC: Functional Suitability
	Short name	User interaction with Virtual assistant
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Virtual Assistant SHOULD interact with the user through a Question-Answer model.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Interaction of the user with the Virtual assistant.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-6
	Dependencies	USR-055, USR-058, USR-059, USR-061, USR-082, USR-118
	Type	FUNC: Functional Suitability
	Short name	User navigation assistance
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Virtual Assistant MUST be able to assist the user navigate into the TRUSTEE dashboard.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The virtual assistant to be able to help the user navigate their way into TRUSTEE.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-7
	Dependencies	USR-059, USR-113
	Type	USE: Usability
	Short name	User input to Robots
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE RPA Robots SHOULD be able to get input from the user.
	Additional Information	None

	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The Robots to be able to stop execution and ask the user to provide information or complete a task.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-8
	Dependencies	USR-048, USR-55, USR-061
	Type	USE: Usability
	Short name	Assist the user in providing data and/or AI/ML models
	Personas	Data Provider, Model Provider
	Description	The TRUSTEE Virtual Assistant SHOULD be able to assist the user to register their data as part of the Distributed Data Sources of TRUSTEE and/or provide their AI/ML models to TRUSTEE.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The user's data registration to TRUSTEE to become easier.

T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-9
	Dependencies	USR-051, USR-52, USR-55, USR-058, USR-061
	Type	USE: Usability
	Short name	Virtual assistant complex queries
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Virtual Assistant SHOULD be able to help users construct complex queries in a visual or text-based way.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Easier complex queries for the user to construct.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-10
	Dependencies	USR-55, USR-056, USR-061, USR-113
	Type	USE: Usability

	Short name	Data consumption automations
	Personas	Consumer
	Description	The TRUSTEE Virtual Assistant SHOULD provide automated procedure regarding data consumption.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Provide automations for data consumption to the user.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-11
	Dependencies	USR-060
	Type	FUNC: Functional Suitability
	Short name	Show data types on searches
	Personas	Consumer
	Description	The TRUSTEE Dashboard SHOULD show the types of data that are available in TRUSTEE's data-catalogs.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation

A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Show data types to the user while searching.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-FUNC-12
	Dependencies	USR-063
	Type	USE: Usability
	Short name	Access tutorials and examples
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Dashboard SHOULD provide the user a way to access the tutorials and example applications.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Easy way for the user to access tutorials and example applications.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-NFUNC-1
	Dependencies	USR-057
	Type	SEC: Security
	Short name	Automation Manager security
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Automation Manager MUST execute processes only from authenticated users.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Make sure the Automation Manager is secure by authentication.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-NFUNC-2
	Dependencies	USR-047, USR-049, USR-050, USR-057, USR-059, USR-062
	Type	USE: Usability
	Short name	User-friendliness
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Dashboard SHOULD display information clearly, consistently, and efficiently, in a user-friendly way.
	Additional Information	None

	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Make sure that TRUSTEE Dashboard is user friendly, consistent and efficient.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-NFUNC-3
	Dependencies	USR-049, USR-050, USR-059, USR-058
	Type	USE: Usability
	Short name	Efficient dashboard
	Personas	Data Provider, Model Provider, Consumer
	Description	The TRUSTEE Dashboard SHOULD make sure that all essential information is immediately accessible.
	Additional Information	None
	Priority (MoSCoW)	S: Should-have. Desirable requirement
M	Measures	
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	Make sure that TRUSTEE Dashboard contains all essential information right up.

T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

	Template Field	Description
S	ID	Req-DA-NFUNC-4
	Dependencies	USR-054, USR-057
	Type	SEC: Security
	Short name	User authentication and registration.
	Personas	Data Provider, Model Provider, Consumer
	Description	The users of the TRUSTEE Platform MUST be able to securely register and authenticate themselves through the dashboard.
	Additional Information	None
	Priority (MoSCoW)	M: Must-have. Mandatory requirement
M	Measures	Implementation and validation
A	Achievable	The combined expertise and technical proficiency of our responsible partner, along with TRUSTEE's application of Agile methodology, ensure the feasibility of our goals. The utilization of the Agile approach facilitates continual risk mitigation and problem resolution throughout the course of development and integration processes.
R	Objectives	The TRUSTEE Dashboard will offer/visualize a way for the user to authenticate and register to TRUSTEE platform, through the use of AM.
T	Timeline	Year 1: independent test of functionality Year 2: verification of integration with other ABBs Year 3: validation against pilot use cases

PART V: TRUSTEE ARCHITECTURE

PART V SUMMARY

This part of the deliverable focuses on the TRUSTEE architecture. It gives insight on a) TRUSTEE architecture development process, framework, modeling language and recommended tool, b) identifies and elaborated TRUSTEE architecture principles, c) presents the TRUSTEE architecture with special focus on TRUSTEE ABBs and their external interfaces and decompositions. The current version of architecture and ABBs is a result of multiple iterations, discussions, and alignments between all WP2 partners and TRUSTEE consortium members and serves as a basis for future work and technical alignment which will be reported in the next iteration of the architecture within deliverable D2.2.

TRUSTEE ARCHITECTURE DEVELOPMENT

The initial version of the TRUSTEE architecture presented in this document is based on the TRUSTEE architecture principles derived from the project's technical description (as presented below in the section TRUSTEE Architecture Principles), as well as the initial discussions among the technical coordinator and WP2 participants. It represents and defines concepts that provide the basis of TRUSTEE functionalities, as stated by the project's technical specifications and outputs.

Since the process of the architecture definition is iterative, the second version of the TRUSTEE architecture will be refined by including additional discussion and alignment results between technical experts contributing to WP2, requirements defined as part of the same work package, and compliance with standards.

Work based on standards improves the common understanding, efficiency, and quality as it minimizes the ambiguities and time it takes until all participants share the same understanding. This is especially important and useful when dealing with diverse participants from different countries and organizations with different backgrounds as is the case in the TRUSTEE project.

Architecture Framework: TOGAF

The Open Group Architecture Framework (TOGAF) is a framework - a detailed method and a set of supporting tools - for developing an enterprise architecture. The TOGAF Standard considers an "enterprise" to be any collection of organizations that have common goals, which we believe is well suited to projects such as TRUSTEE where the TRUSTEE Consortium can be considered an enterprise in this sense.

Consulted is the TOGAF® Standard, 10th Edition <https://www.opengroup.org/togaf/10thedition>, freely available for viewing online and using by any organization⁹⁹³ (with restrictions on commercial use).

TOGAF supports common architecture domains including business architecture, data architecture, application architecture, and technology architecture. TOGAF Architecture Development Method (ADM) provides a tested and repeatable process for developing architectures and is illustrated in the following figure (Figure 67). More information on the TOGAF ADM, as the core of the TOGAF Standard, can be found on <https://pubs.opengroup.org/togaf-standard/adm/index.html>.

⁹⁹³ https://pubs.opengroup.org/togaf-standard/index.html#tag_01_03_01 (accessed 5 January 2023)

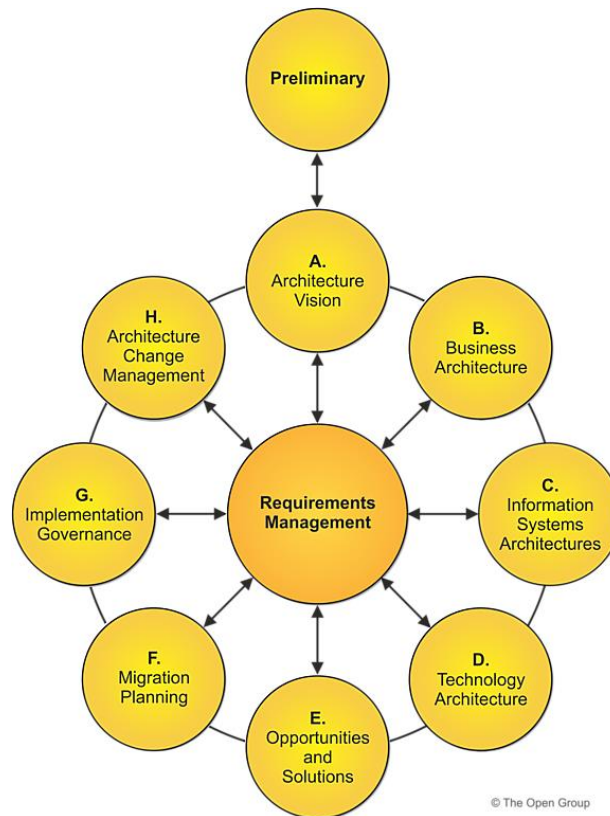


Figure 67: TOGAF Architecture Development Cycle (image from ⁹⁹⁴)

The flexibility of the TOGAF framework is utilized in the specific TRUSTEE architectural style by combining TOGAF features most useful for TRUSTEE.

Each of the TOGAF Architecture Development Cycle phases was addressed, although in the scope and details determined by the TRUSTEE WP2 participants to develop TRUSTEE architecture. More specifically the phases and related activities were:

- 1) Preliminary, initiation of project and work, identification of context, detailed consideration of TRUSTEE Description of Work, identification of stakeholders, definition of TRUSTEE manifesto, initiation of work on continuously updated Glossary, identification of Architecture Principles;
- 2) Architecture Vision, where the business goals and drivers are defined;
- 3) Business Architecture, where the TRUSTEE structure, processes, and capabilities are analyzed;
- 4) Information Systems Architecture, where the technology infrastructure and information requirements are preliminarily addressed;
- 5) Technology Architecture, where the specific technologies and platforms are selected. However, the work initiated within WP2 and this deliverable focused on preconditions, such as review of the state of the art, identification of different User-centred, Legal, Socio-ethical, and Technical requirements, architecture design with identification of main architecture building blocks, their capabilities and functionalities. The selection of specific technologies and

⁹⁹⁴ <https://pubs.opengroup.org/togaf-standard/introduction/chap03.html>

frameworks will be done in implementation work packages (primarily WP3,WP4), although some insights and recommendations can be found in this deliverable.

6) Opportunities and Solutions, where potential projects and solutions are identified. In TRUSTEE, the DoW prescribes already the three iterations of the deliverable, while continuous discussions were done not only within WP2 but also cross-WP (after the implementation WPs have started, that is M6).

7) Migration Planning, where a roadmap for implementing the architecture is created. Here the business value to each of the work packages was confirmed, as this was given in TRUSTEE DoW.

8) Implementation Governance, where the architecture is monitored and controlled during implementation. The WP2 Leaders Task Force (WP2LTF) was established at the beginning of the WP2 under the coordination of WP2 leader, comprising of: WP2 leader, all WP2 task leaders (T2.1, T2.2., T2.3, T2.4, T2.5), project Technical Management team, as well as WP3 and WP4 leaders and legal experts, together with their deputies in case of their absence, on regular biweekly telcos.

9) Architecture Change Management, where changes to the architecture are managed, under the coordination of Technical Management team, involving ABB leaders and WP2 leader. When some change was agreed it was propagated across all ABBs to be aligned. Selected Architecture Decisions were documented and propagated to all aforementioned parties.

SCOPE OF THE TRUSTEE ARCHITECTURE

The four dimensions that typically define and limit the scope of an architecture, and their mapping to TRUSTEE are as follows:

- I. Breadth – TRUSTEE is a project running for 42 months from July 2022 until its expected end in 2025, between 22 partners. The main effort of Architecture work is confined within WP2 and more concretely in this deliverable and its respective upcoming versions, namely D2.2 and D2.3.
- II. Depth – to ensure the proper level of detail is captured for TRUSTEE Architecture, the WP2LTF was established by the WP2 leader with the main leaders of all the WP2 tasks in addition to leaders of the technical implementation WPs, namely WP3 and WP4, under the supervision of the project Technical Manager. Continuous discussions and alignments were done to ensure the proper outcomes are achieved and delivered as sufficient for its purpose, primarily common understanding within the project, needed communication with the TRUSTEE relevant external entities, and technical implementation work packages realizing the TRUSTEE.
- III. Time period – as the vision of TRUSTEE is becoming more transparent, the new agreed decisions are made considering all the inputs coming from the TRUSTEE partners with multidisciplinary and diverse background. TRUSTEE architecture work is thus evolutionary or incremental and is a result of many discussions, alignments, and periodic updates. The project schedule and the deadlines, especially project milestones and implementation plan, under the tight control from the project Technical Management Team and ultimately Project Management Board, ensures that the syndrome of “moving target” is avoided and that the best possible balance in terms of the scope and depth is achieved.
- IV. Architecture domains – the main TRUSTEE User groups/Personas (in our case primarily the Data providers, Model Providers, Consumers, and Developers) need to understand how TRUSTEE addresses their needs and which benefits they can expect. For this reason, the *business domain* side of the TRUSTEE offering is extremely important, since the critical step is to ensure the wide use of the TRUSTEE Platform to create more offering and thus better added value for all parties. The *data, application and technology domain* are needed mostly for the implementation purposes, as well as for

in-depth understanding of the TRUSTEE innerworkings and realization in terms of implementation and deployment.

RECOMMENDED MODELLING LANGUAGE: ARCHIMATE

To be aligned as much as possible, the WP2 leadership proposed the Archimate modeling language to be used for the description of the TRUSTEE architecture, however, and in order to be inclusive to some degree, alternative approaches were also allowed from all the partners working on specific ABBs.

ArchiMate is a modelling language for enterprise architecture based on the ISO/IEC/IEEE 42010:2022 Software, systems and enterprise — Architecture description⁹⁹⁵ (previously IEEE 1471) standard. It provides a standardized vocabulary and notation for describing the components and relationships of an enterprise architecture, and is intended to help organizations understand, analyze, and communicate their complex business and IT landscapes. ArchiMate is developed and maintained by The Open Group, an international consortium of companies and organizations that promotes open standards and best practices in the field of information technology.

The ArchiMate® Specification, a standard of The Open Group, is an open and independent modelling language for Enterprise Architecture. It provides instruments to enable Enterprise Architects to describe, analyze, and visualize the relationships among business domains in an unambiguous way. It defines a common language for describing the construction and operation of business processes, organizational structures, information flows, IT systems, and technical infrastructure.

The standard provides a set of entities and relationships with their corresponding iconography for the representation of Architecture Descriptions. The ArchiMate ecosystem also supports an exchange format in XML which allows model and diagram exchange between tools.

ArchiMate® 3.1 Specification, *a Standard of The Open Group can be accessed on this web link:*
<https://pubs.opengroup.org/architecture/ArchiMate3-doc/>

The following figure outlines the top-level hierarchical structure of the language:

- A model is a collection of (abstract) concepts – a concept is either an element or a relationship
- An element is either a behavior element, a structure element, a motivation element, or a composite element

⁹⁹⁵ <https://www.iso.org/standard/74393.html>

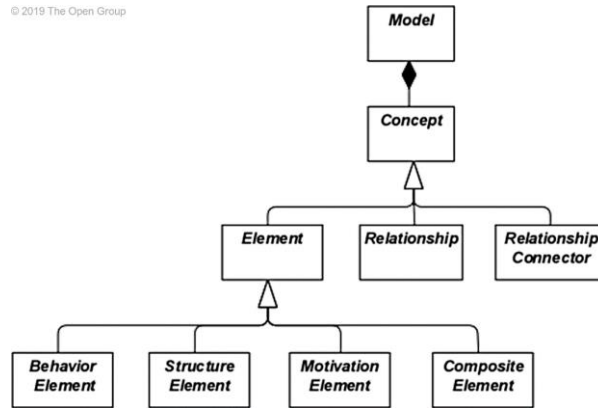


Figure 68: Top-Level Hierarchy of ArchiMate Concepts (source: <https://pubs.opengroup.org/architecture/ArchiMate3-doc/chap03.html>)

ARCHIMATE FULL FRAMEWORK

ArchiMate framework, as visualized in the following figure (Figure 69), proposes **layers** and **aspects** as the two main dimensions for organizing all the elements.

Core layers represent the three levels: Business, Application, and Technology, each with its own set of generic elements and relationships that can be used to model various aspects of the enterprise.

- *Business Layer* depicts business services offered to customers,
- *Application Layer* depicts application services that support the business, and the applications that realize them,
- *Technology Layer* depicts technology services such as processing, storage, and communication services needed to run the applications, and the computer and communication hardware and system software that realize those services. Physical elements are included for modelling physical equipment,
- while it defines additional layers for strategy, physical and implementation/migration elements.
- Aspects refer to:
 - i. the active structure (elements representing who/what makes the things),
 - ii. behavior (elements indicating what is made and how it is made), and
 - iii. passive structure (things on which behavior is performed).
 - iv. It also defines a fourth additional aspect with motivational elements (why things are made). Motivation elements can be used to model the reasons and drivers behind the design or change of

	Passive Structure	Behavior	Active Structure	Motivation
Strategy				Layers
Business				
Application				
Technology				
Physical				
Implementation & Migration				

Aspects

ARCHIMATE'S BENEFITS

ArchiMate's benefits include⁹⁹⁶:

-
- ARCHIMATE USER COMMUNITY

ARCHIMATE NOTATION OVERVIEW

TRUSTEE

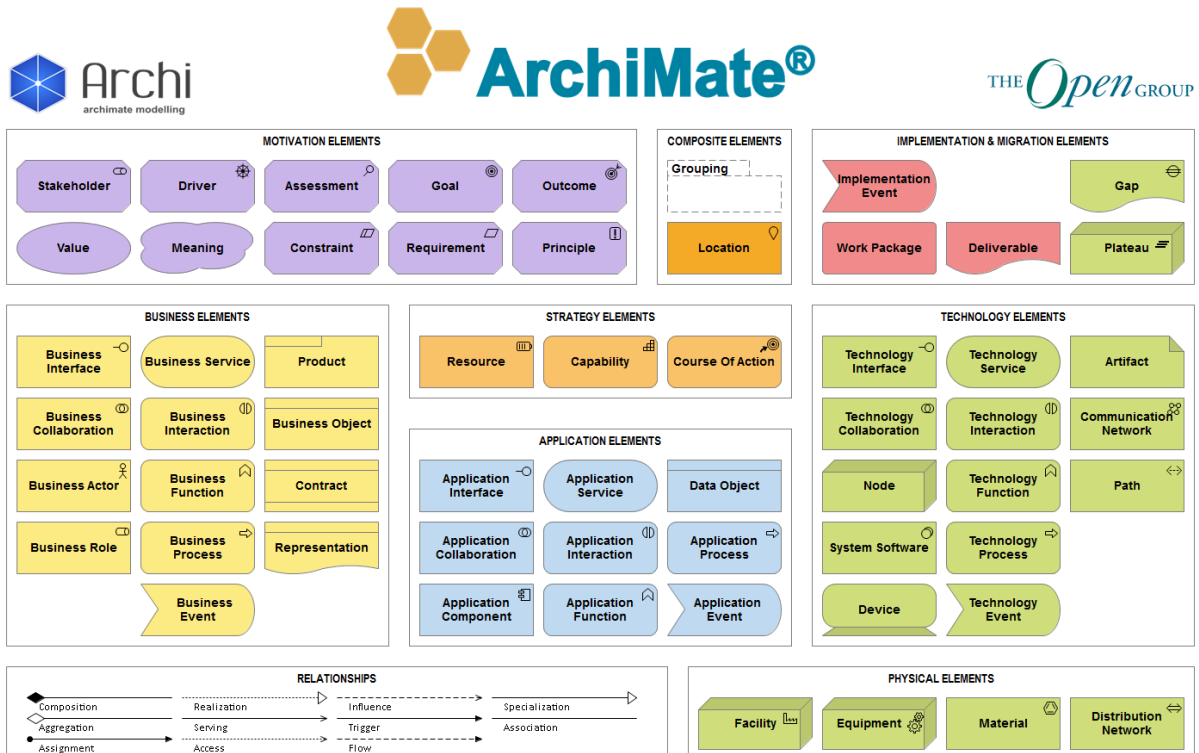


Figure 70: ArchiMate Notation overview (image from⁹⁹⁷)

ARCHIMATE VS UML

Both Archimate and UML, although not exclusively, were used in TRUSTEE, so here we summarize their comparison.

Both ArchiMate and UML are useful modelling languages, but they are used in different contexts and for different purposes. ArchiMate focuses on enterprise architecture, while UML is a more general modelling language for software systems

- ArchiMate is a modelling language designed specifically for enterprise architecture, it provides a set of concepts and notations for describing the components, relationships, and behaviors of an organization's architecture, including business processes, applications, data, and technology infrastructure. ArchiMate is commonly used by enterprise architects to document and analyze the current state of their organization and plan for future changes.
- UML, on the other hand, is a universal modelling language used to model a wide range of systems, not just enterprise architectures. UML provides a set of concepts and notations for describing the structure, behavior, and interaction of software systems. UML diagrams are often used by software developers to design and communicate the structure and behavior of software systems.

⁹⁹⁷ <https://www.archimatetool.com/wp-content/uploads/2018/11/ArchiMate-3.0-Notation-Overview-canvas-Low-saturation-color-scheme.png> , Copyright © 2013-2023 Phillip Beauvoir & Jean-Baptiste Sarrodie. All Rights Reserved.

Both ArchiMate and UML provide graphical notations for modelling but differ in the types of diagrams used and the concepts they represent. For example, ArchiMate diagrams contain concepts such as business processes, actors, and stakeholders, while UML diagrams contain concepts such as classes, objects, and messages.

Mapping of ArchiMate and UML is proposed in following paper: Geicke, T. "ArchiMate to UML mapping." *Adocus AB* (2016)⁹⁹⁸ for the ones who are interested in a more detailed comparison.

TOGAF AND ARCHIMATE SYNERGY

The Open Group created and maintains both TOGAF and the ArchiMate modelling language.

TOGAF and ArchiMate are related but distinct concepts. **TOGAF is a framework for developing (enterprise) architecture, while ArchiMate is a modelling language** for representing the components and relationships of an enterprise architecture. TOGAF provides a high-level structure and methodology for designing, planning, and implementing enterprise IT systems, and includes guidance on how to use different modelling languages, including ArchiMate, to support these activities. In this sense, ArchiMate is a tool that can be used within the TOGAF framework to help organizations create and manage their enterprise architectures.

RECOMMENDED TOOL FOR DESCRIBING TRUSTEE ARCHITECTURE: ARCHI

Similarly, as for the modeling language, a recommendation was also given on the tool to be used for the de, however it was not prescribed as mandatory to use only one tool for describing all ABBs.

Archi (<https://www.ArchiMatetool.com/>) is a free, open-source (MIT license) tool for creating and managing architectural diagrams. It is based on the ArchiMate modelling language, and is designed to be intuitive and easy to use for architects, business analysts, and other professionals who need to visualize and document the structure and behavior of complex systems. Archi allows users to create diagrams using a wide range of predefined elements and relationships, and to customize and extend the modelling language as needed to capture the specific details of their architecture. It also provides a range of features for collaboration, version control, and publishing, making it a powerful tool for supporting the entire enterprise architecture process.

From <https://www.archimatetool.com/>: *Archi® fulfils the needs of most Enterprise Architects and associated stakeholders, and has been designed to elegantly provide the main features required for ArchiMate modelling and is used globally by banks, insurance companies, industry, EA consultants, training organisations, universities, and students. It is probably the most popular ArchiMate modelling tool and is downloaded on average about 6,000 times every month.*

⁹⁹⁸ <https://www.adocus.com/media/21703/archimate-to-uml-mapping-whitepaper.pdf> (23.2.2023.)

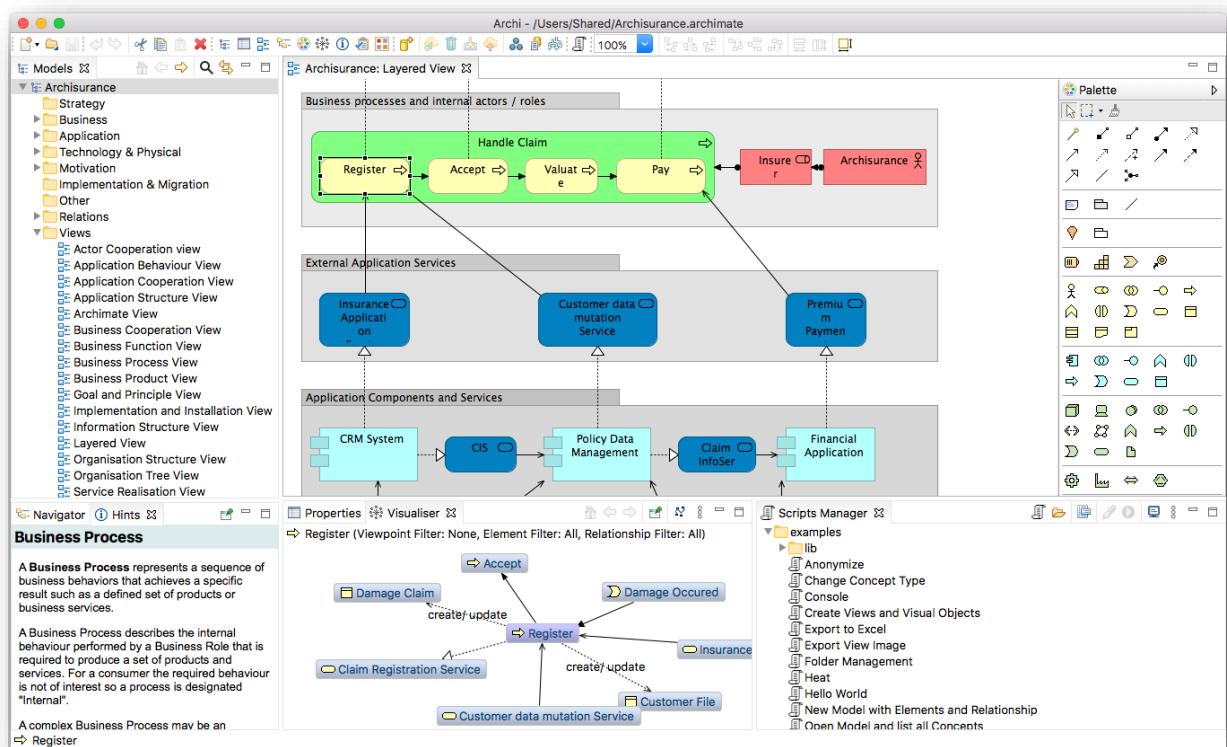


Figure 71: Archi tool IDE (image from⁹⁹⁹)

At the time of the beginning of work in D2.1, version of the Archi tool is v4.10, released in September 15 2022 with support of the latest ArchiMate® 3.1 specification, while current (in May 2023) the latest version is 5.0.2, released in April 2023, with support for the ArchiMate® 3.2 specification.

More on Archi is available in User Guide (Version 5.0.2 has 161 pages) available here: <https://www.ArchiMatetool.com/downloads/archi/Archi%20User%20Guide.pdf>

TRUSTEE ARCHITECTURE PRINCIPLES

Principles driving the work of the TRUSTEE architecture are summarized in the earliest TRUSTEE internal but also external communication with the origin in WP2 and the work reported in this deliverable. In this chapter the principles are elaborated in more detail, starting with the explanation of the template (or format) used for their description.

TRUSTEE architecture principles reflect a level of consensus across the project and are used to govern the architecture process, affecting the development, maintenance, and use of the TRUSTEE Architecture.

TEMPLATE FOR DESCRIBING ARCHITECTURAL PRINCIPLES

⁹⁹⁹ <https://www.archimatetool.com/>

Template for describing the TRUSTEE architecture principles (a set of principles that relate to architecture work) is inspired by and compliant with one from the TOGAF.

NAME	Represents the rule's essence in short. Should not include any specific technical terms.
Statement	Should succinctly and unambiguously communicate the fundamental rule. For the most part, the principles statements for managing information are similar from one organization to the next. It is vital that the principles statement is unambiguous.
Rationale	Should convey the fundamental rule in a clear and concise manner. The principles statements for managing information are generally the same across organizations. It is essential that the statement of the principles are clear.
Implications	Should emphasize the resources, costs, and activities needed to implement the principle, as well as the requirements. The impact to the project and consequences of adoption should be clearly stated.

FRAMEWORK FOR SECURE AND TRUSTWORTHY DATA FEDERATION

NAME	Framework for secure and trustworthy data federation
Statement	TRUSTEE will be used as an enabler for secure and trustworthy data federation between different Open Data Spaces
Rationale	Considering identification, authentication and authorization mechanisms, and established mechanisms on EU level TRUSTEE will facilitate the vision of Open Data Spaces and deliver tools and practices for enabling data federation in an encrypted domain.
Implications	TRUSTEE will facilitate the federation of Data Spaces TRUSTEE will not cater only to a specific domain or sector

Leveraging EU Initiatives and Existing Technologies and Standards

NAME	Leveraging EU initiatives and existing technologies and standards
Statement	TRUSTEE will utilize existing and well-established standards, technologies, and initiatives.
Rationale	TRUSTEE will build on the existing work and know how to deliver new added value. Solution development should be driven by the TRUSTEE needs. Standards help ensure consistency, thus improving the ability to manage systems and improve user satisfaction
Implications	Interoperability standards and industry standards will be followed unless there is a compelling business reason to implement a non-standard solution

DATA IS AN ASSET

NAME	Data is an asset
Statement	Data is an asset that has value to the individual or organization and as such should be managed accordingly.
Rationale	Data is a valuable resource. It is the foundation for decision-making, so we must treat it with care, to properly store it, curate it and guard it.
Implications	Access to data does not grant rights to modify or disclose the data.

DATA IS SECURED

NAME	Data is secured
Statement	Data is protected from unauthorized use and disclosure
Rationale	Release of information via relevant legislation must be balanced against the need to restrict the availability of classified, proprietary, and sensitive information. Existing laws and regulations require the safeguarding of privacy of data while permitting free and open access. Access to data is authenticated and authorized and while data reside on the premise of Data Providers.
Implications	To adequately provide access to open information while maintaining secure information, security needs must be identified and developed at the data level. Data requires "de-confidentiality" analyses and procedures to maintain proper control. Data security safeguards can be put in place to restrict access to "view only" or "never see". To provide appropriate access to open information yet maintain security, the security restrictions must be identified and implemented at the data level, not at the application level. Security must be designed into data elements from the beginning; it cannot be added later. In TRUSTEE homomorphic encryption is used to prevent undesired information interception and confidentiality loss, and to enable third parties to apply operations on encrypted data without revealing the original data values.

DATA SHOULD BE FAIR

NAME	Data should be FAIR
Statement	Metadata and data should follow FAIR principles (https://www.go-fair.org/fair-principles/) to be easy to find and access for both humans and computers. Metadata and data should be well-described using broadly applicable language and vocabularies.
Rationale	There is an evident need to improve the infrastructure supporting the reuse of different data.
Implications	TRUSTEE shall support the FAIR principles.

APPLICATIONS ARE TECHNOLOGY INDEPENDENT

NAME	Applications are technology independent
Statement	Applications (that will utilize the data arriving via TRUSTEE) are independent of specific technology choices and therefore can operate on a variety of technology platforms
Rationale	Independence of applications from the underlying technology allows applications to be developed, upgraded, and operated in the most cost-effective and timely way
Implications	Middleware should be used to decouple applications from specific software solutions

TRUSTEE ARCHITECTURE DESCRIPTION

LEGEND: ARCHIMATE CORE ELEMENTS AND RELATIONSHIPS

For better understanding of the following figures please consider previous chapter ArchiMate Full Framework. In addition, the following tables summarize elements and relationship notations used in the subsequent chapters.

Table 0.15 ArchiMate Core Elements

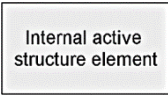
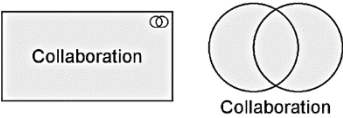
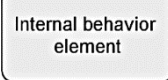
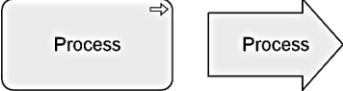
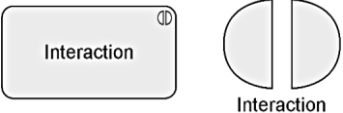
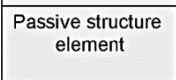
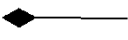
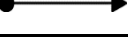
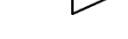
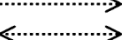
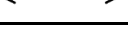
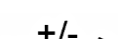
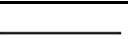
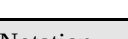
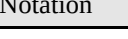
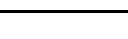
Element	Specializations	Definition	Notation
Active Structure			
Internal active structure element		Represents an entity that is capable of performing behavior.	
	Collaboration	Represents an aggregate of two or more internal active structure elements, working together to perform some collective behavior.	
Interface (external active structure element)		Represents a point of access where one or more services are exposed to the environment.	
Behavior			
Internal behavior element		Represents a unit of activity that can be performed by one or more active structure elements.	
	Process	Represents a sequence of behaviors that achieves a specific result.	
	Function	Represents a collection of behavior based on specific criteria, such as required resources, competencies, or location.	
	Interaction	Represents a unit of collective behavior that must be performed by two or more internal active structure elements, either assigned directly or aggregated in a collaboration.	
Service (external behavior element)		Represents an explicitly defined exposed behavior.	
Event		Represents a state change.	
Passive Structure			
Passive structure element		Represents an element on which behavior is performed.	

Table 0.16 ArchiMate Relationships

Structural Relationships	Notation	Role Names

Composition	Represents that an element consists of one or more other concepts.		← composed of → composed in
Aggregation	Represents that an element combines one or more other concepts.		← aggregates → aggregated in
Assignment	Represents the allocation of responsibility, performance of behavior, storage, or execution.		← assigned to → has assigned
Realization	Represents that an entity plays a critical role in the creation, achievement, sustenance, or operation of a more abstract entity.		← realizes → realized by
Dependency Relationships		Notation	Role Names
Serving	Represents that an element provides its functionality to another element.		← serves → served by
Access	Represents the ability of behavior and active structure elements to observe or act upon passive structure elements.	  	← accesses → accessed by
Influence	Represents that an element affects the implementation or achievement of some motivation element.		← influences → influenced by
Association	Represents an unspecified relationship, or one that is not represented by another ArchiMate relationship.	 	associated with ← associated to → associated from
Dynamic Relationships		Notation	Role Names
Triggering	Represents a temporal or causal relationship between elements.		← triggers → triggered by
Flow	Represents transfer from one element to another.		← flows to → flows from
Other Relationships		Notation	Role Names
Specialization	Represents that an element is a particular kind of another element.		← specializes → specialized by
Relationship Connectors		Notation	Role Names
Junction	Used to connect relationships of the same type.	 (And) Junction  Or Junction	

STRATEGY VIEWS

Strategy typically expresses how the organization (in our case TRUSTEE) wants to create value for its stakeholders, capabilities it needs for that, and resources needed to support these capabilities¹⁰⁰⁰.

¹⁰⁰⁰ ArchiMate® 3.1 Specification, Copyright © 2012-2019 The Open Group, Strategy Elements
<https://pubs.opengroup.org/architecture/archimate3-doc/chap07.html#>

The Strategy View is one of the core views in ArchiMate, and it provides a high-level view of an organization's, in our case TRUSTEE consortium, business strategy.

CAPABILITIES AND RESOURCES MAP

Following figure (Figure 72) shows the TRUSTEE's capabilities and resources. Capability represents an ability that the TRUSTEE consortium and the envisioned TRUSTEE Federated Platform possess. Resources stand for assets owned or controlled by a TRUSTEE consortium.

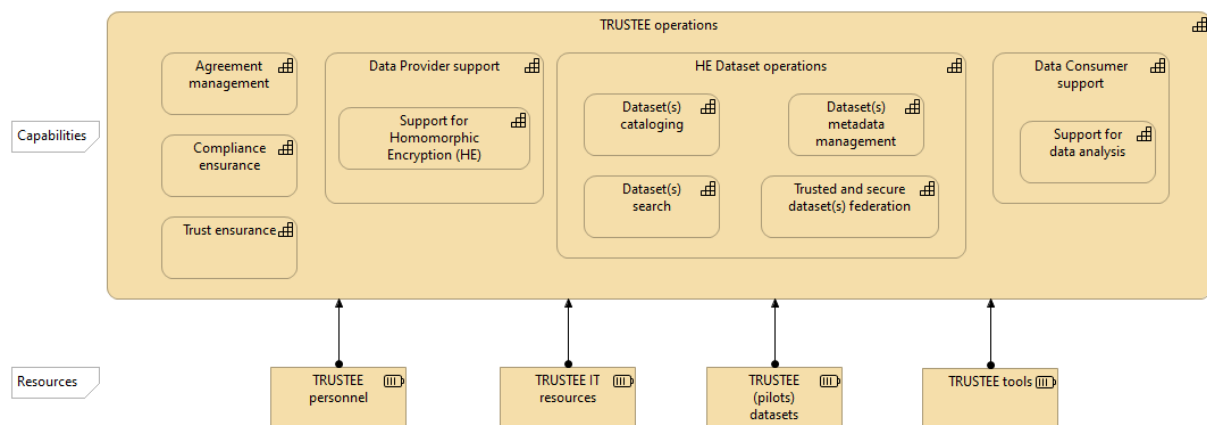


Figure 72: TRUSTEE capabilities and resources

BUSINESS STRATEGY VIEW

Following figure (Figure 73) shows TRUSTEE business strategy diagram in ArchiMate to help business leaders and other stakeholders to understand the strategic vision of the TRUSTEE and how it will be executed.

The figure depicts connections between TRUSTEE stakeholders, drivers, goals, expected outcomes, courses of action, business capabilities and resources. TRUSTEE stakeholders and drivers are motivating TRUSTEE to pursue a particular course of action, the goals that the TRUSTEE hopes to achieve, the expected outcomes of those goals, and the (business) capabilities and resources that will be used to carry out the plan.

- Stakeholders: the organizations or individuals who have an interest in t TRUSTEE and its activities.
- Drivers: the factors that are motivating TRUSTEE to take action.
- Goals: the desired outcomes that TRUSTEE is trying to achieve.
- Expected outcomes: the specific results that TRUSTEE hopes to achieve by pursuing its goals.
- Courses of action: the specific steps that TRUSTEE plans to take in order to achieve its goals.
- Business capabilities: the abilities and resources that TRUSTEE has to carry out its courses of action.
- Resources: the tangible and intangible assets that the business has at its disposal. Examples might include financial resources, physical assets like buildings or equipment, and intellectual property.

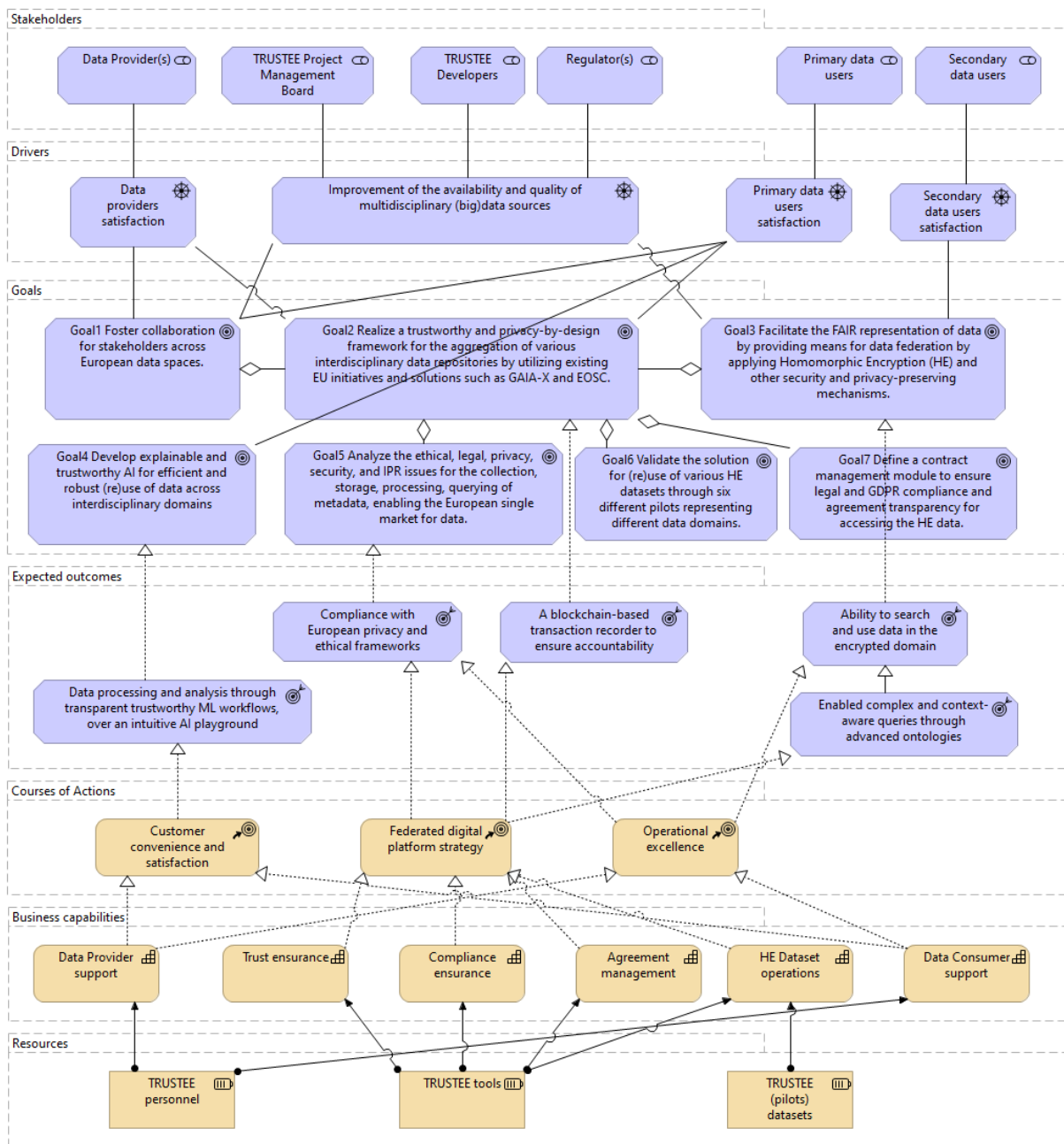


Figure 73: TRUSTEE stakeholders, drivers, goals, expected outcomes, courses of action, business capabilities and resources

BUSINESS VIEWS

In the ArchiMate modelling language, "Business Views" represent high-level views of an organization's business processes, capabilities, and organizational structure. TRUSTEE Business Views provide a holistic understanding of TRUSTEE business operations and are used to model and analyze TRUSTEE strategic goals and objectives.

BUSINESS LAYERS VIEW

Business layer elements model the operational organization of an enterprise (in our case TRUSTEE) in a technology-independent manner¹⁰⁰¹.

- TRUSTEE (Business) Actors are (business) entities that can perform behavior.
- TRUSTEE (Business) Services are services (as explicitly defined behaviors) offered by TRUSTEE to its customers (primarily Data Providers, Model Providers, and Consumers) or other stakeholders.
- TRUSTEE Business Interface represents a point of access where a business service is made available to the environment.
- TRUSTEE Business Processes are activities and workflows that the TRUSTEE consortium performs to deliver its services.

Following figure (Figure 74) shows TRUSTEE end-user groups, business services, business interfaces (channels) and business processes.

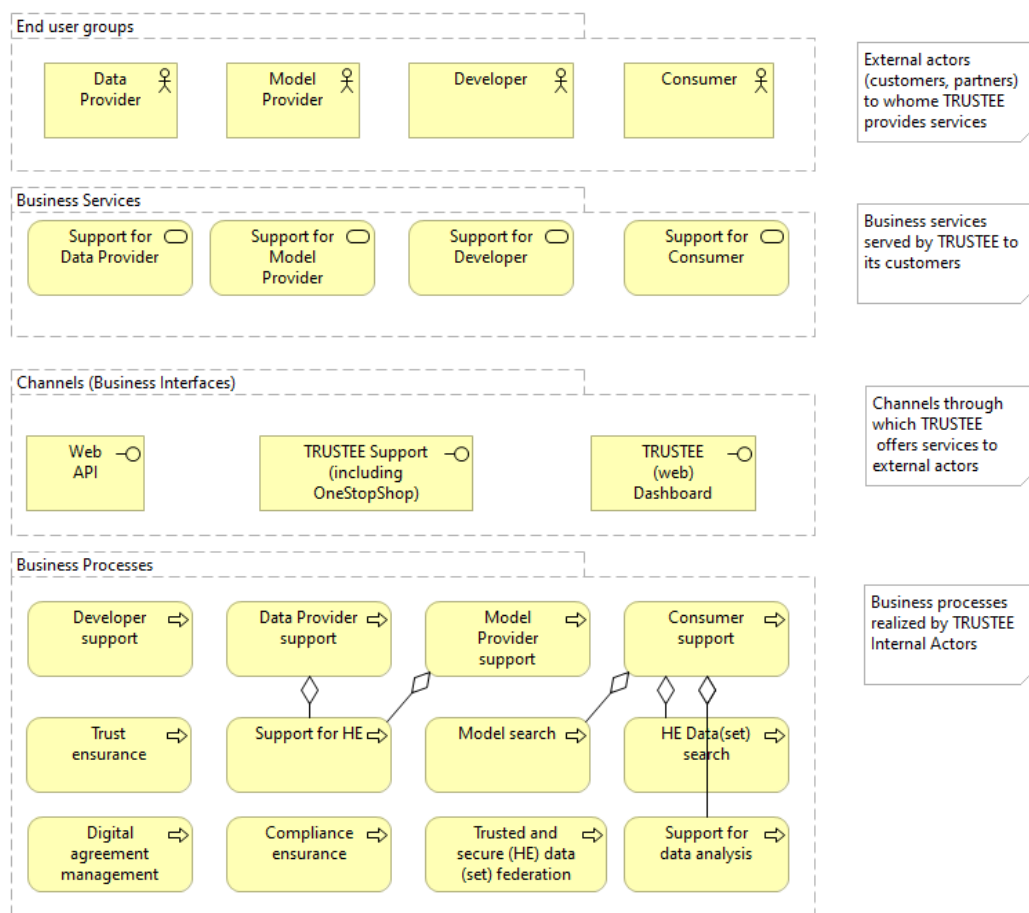


Figure 74: TRUSTEE end-user groups, business services, interfaces, processes

¹⁰⁰¹ ArchiMate® 3.1 Specification, Copyright © 2012-2019 The Open Group, Business Layer <https://pubs.opengroup.org/architecture/archimate3-doc/chap08.html#>

CAPABILITIES TO BUSINESS PROCESSES DIAGRAM

TRUSTEE Business Process diagram shows the processes and activities that the TRUSTEE consortium performs to deliver value to its customers (or users) and how they are related to the TRUSTEE consortium capabilities.

Following figure (Figure 75) shows connections between TRUSTEE business capabilities and business processes.

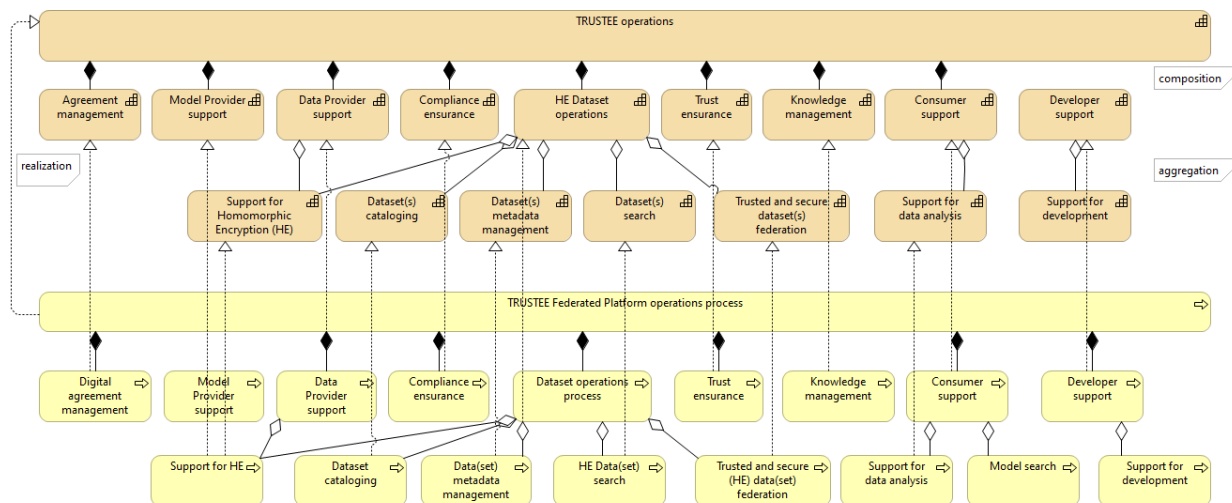


Figure 75: TRUSTEE business capabilities to business processes

DATA ACTIVITY MAP

TRUSTEE Dataset activity map, shown in following figure (Figure 76), provides an overview of the TRUSTEE Business Services and Business Processes related to the Data(sets) in TRUSTEE.

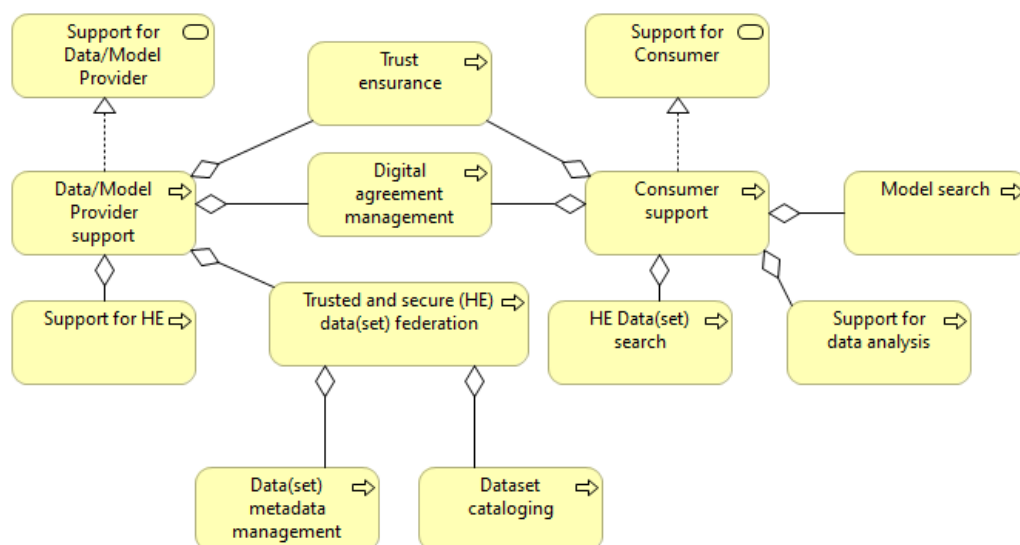


Figure 76: Data(set) activities in TRUSTEE

DATA PROVIDER DIAGRAM

Following figure (Figure 77) shows how Data Provider business service is mapped to business processes and further to application services, processes, and data.

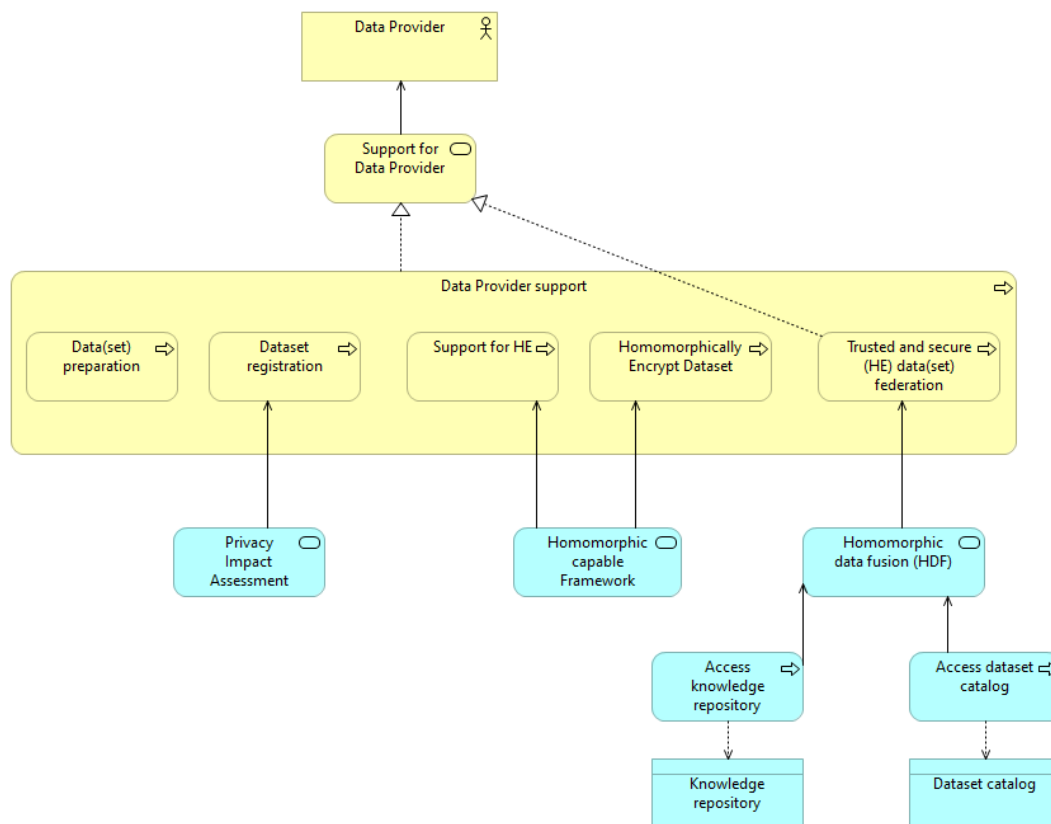


Figure 77: Data Provider business service to business processes to application services, processes, and data

MODEL PROVIDER DIAGRAM

Following figure (Figure 78) shows how Model Provider business service is mapped to business processes and further to application services, processes, and data.

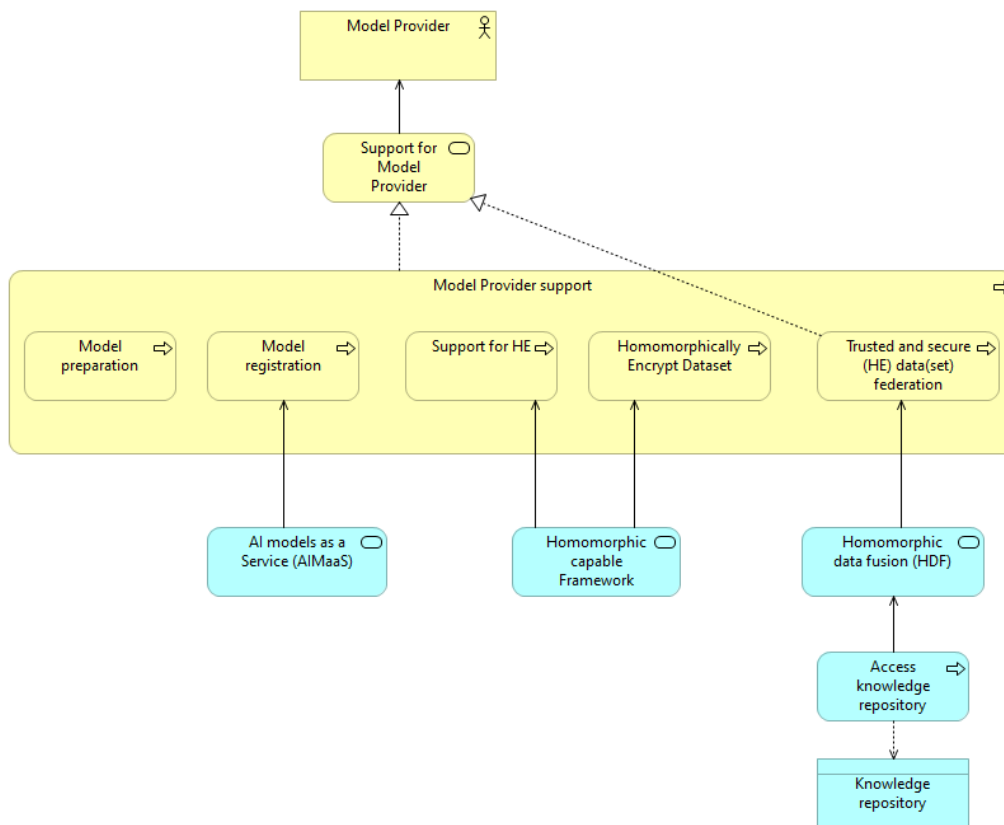


Figure 78: Model Provider business service to business processes to application services, processes, and data

CONSUMER DIAGRAMS

The following figure (Figure 79) shows how Consumer business service is mapped to business processes and further to application services, processes, and data when Consumer is consuming data.

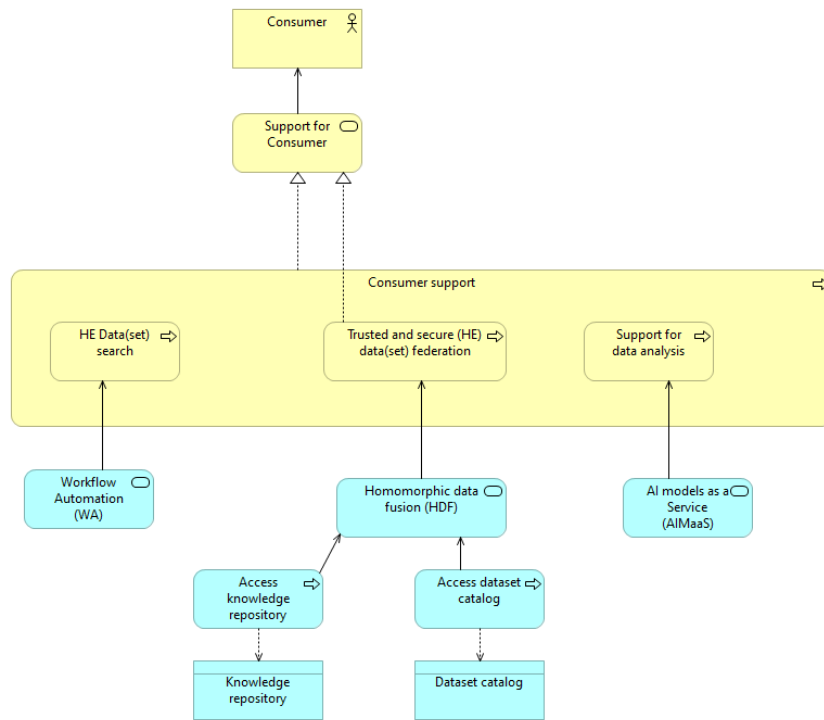


Figure 79: Consumer using data - business service to business processes to application services, processes, and data

The following figure (Figure 80) shows how Consumer business service is mapped to business processes and further to application services, processes, and data when Consumer is consuming models.

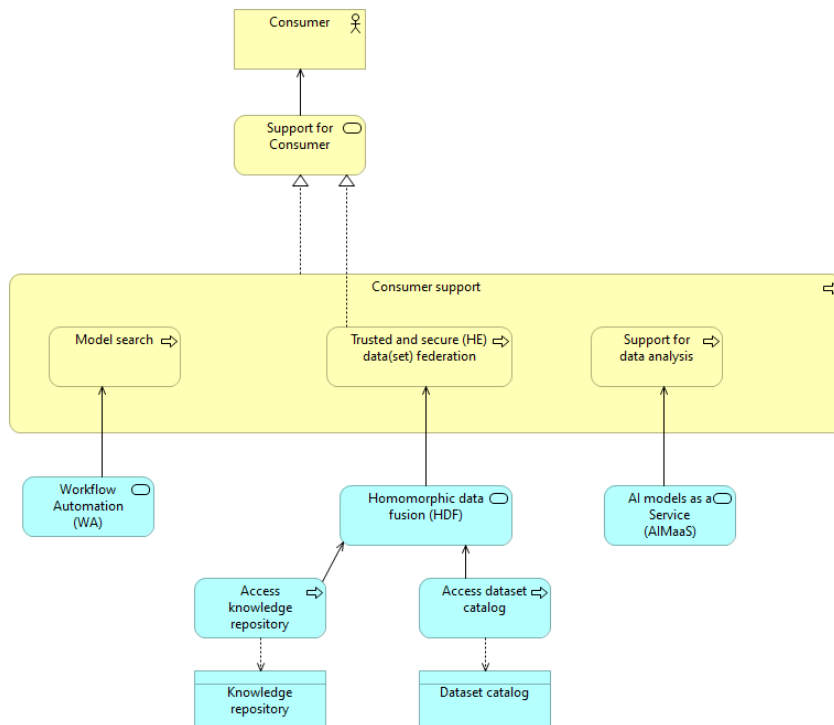


Figure 80: Consumer using model - business service to business processes to application services, processes, and data

DEVELOPER DIAGRAM

The following figure (Figure 81) shows how Developer business service is mapped to business processes and further to application services, processes, and data.

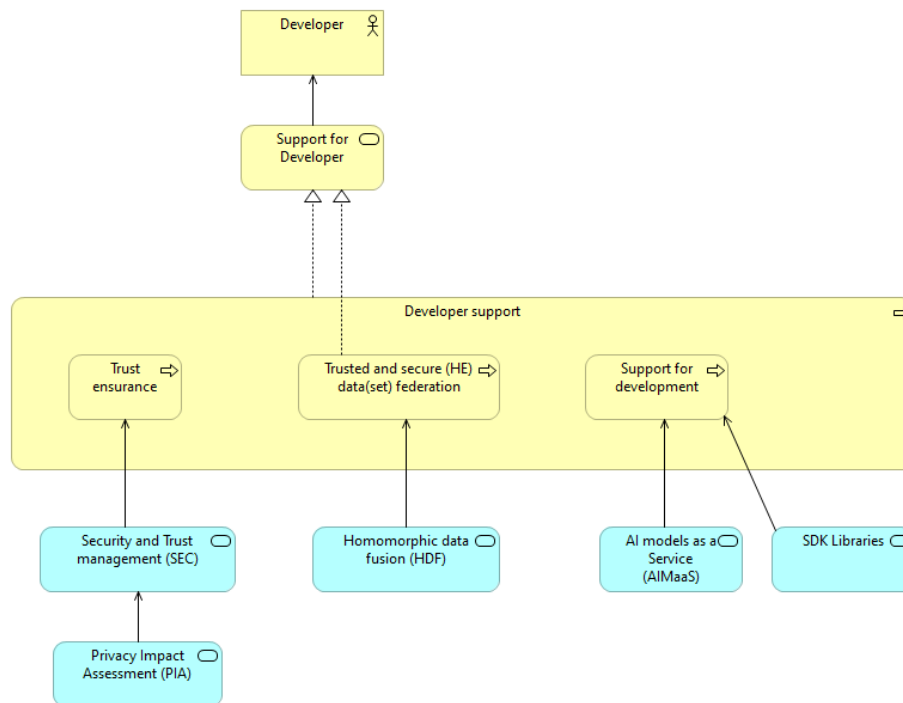


Figure 81: Developer business service to business processes to application services, processes, and data

APPLICATION VIEWS

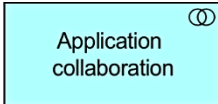
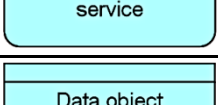
Application layer elements model the Application Architecture that describes the structure, behavior, and interaction of the applications of the enterprise (in our case TRUSTEE).

ELEMENT DEFINITION AND NOTATION

The following table summarizes the Elements definition and notation adopted from ArchiMate® 3.1 Specification ¹⁰⁰².

¹⁰⁰² ArchiMate® 3.1 Specification, Copyright © 2012-2019 The Open Group, Application Layer <https://pubs.opengroup.org/architecture/archimate3-doc/chap09.html#>

Table 0.17 Application Layer Elements (from ArchiMate® 3.1 Specification¹⁰⁰³)

Element	Definition	Notation
Application component	Represents an encapsulation of application functionality aligned to implementation structure, which is modular and replaceable.	
Application collaboration	Represents an aggregate of two or more application internal active structure elements that work together to perform collective application behavior.	
Application interface	Represents a point of access where application services are made available to a user, another application component, or a node.	
Application function	Represents automated behavior that can be performed by an application component.	
Application interaction	Represents a unit of collective application behavior performed by (a collaboration of) two or more application components.	
Application process	Represents a sequence of application behaviors that achieves a specific result.	
Application event	Represents an application state change.	
Application service	Represents an explicitly defined exposed application behavior.	
Data object	Represents data structured for automated processing.	

ARCHITECTURE BUILDING BLOCKS (ABB) OVERALL INTERACTION DIAGRAM

Figure 82 shows ABBs interaction diagram illustrating data flows between ABBs at high-level. This type of diagram is frequently referred to also as communication diagram and is used to show the communication paths and relationships between high level building blocks focusing on the relationships between components rather than their individual behavior. The approach towards the definition of this diagram led also to an initial identification of the external communication interfaces between the various architecture blocks of the TRUSTEE Platform. The

¹⁰⁰³ <https://pubs.opengroup.org/architecture/archimate3-doc/chap09.html#> Copyright © 2012-2019 The Open Group, All Rights Reserved, ArchiMate is a registered trademark of The Open Group.

interfaces will be further described in the following sections of this deliverable and will be enhanced and updated, if needed, in future versions of this deliverable, namely: D2.2 and D2.3.

While an overview of all ABBs will be provided in the following sections, including description, sequence diagrams, and external communication interfaces definition, we need to clarify at this point that the Distributed Data Sources, depicted in the Architecture diagram below, refer to the datasets of the Data Providers of TRUSTEE. These data sources are external to TRUSTEE and distributed, as they always reside at the Data Providers' premises and they are not considered an architecture block of the TRUSTEE Platform, rather an extension of the TRUSTEE Data Lake (DL).

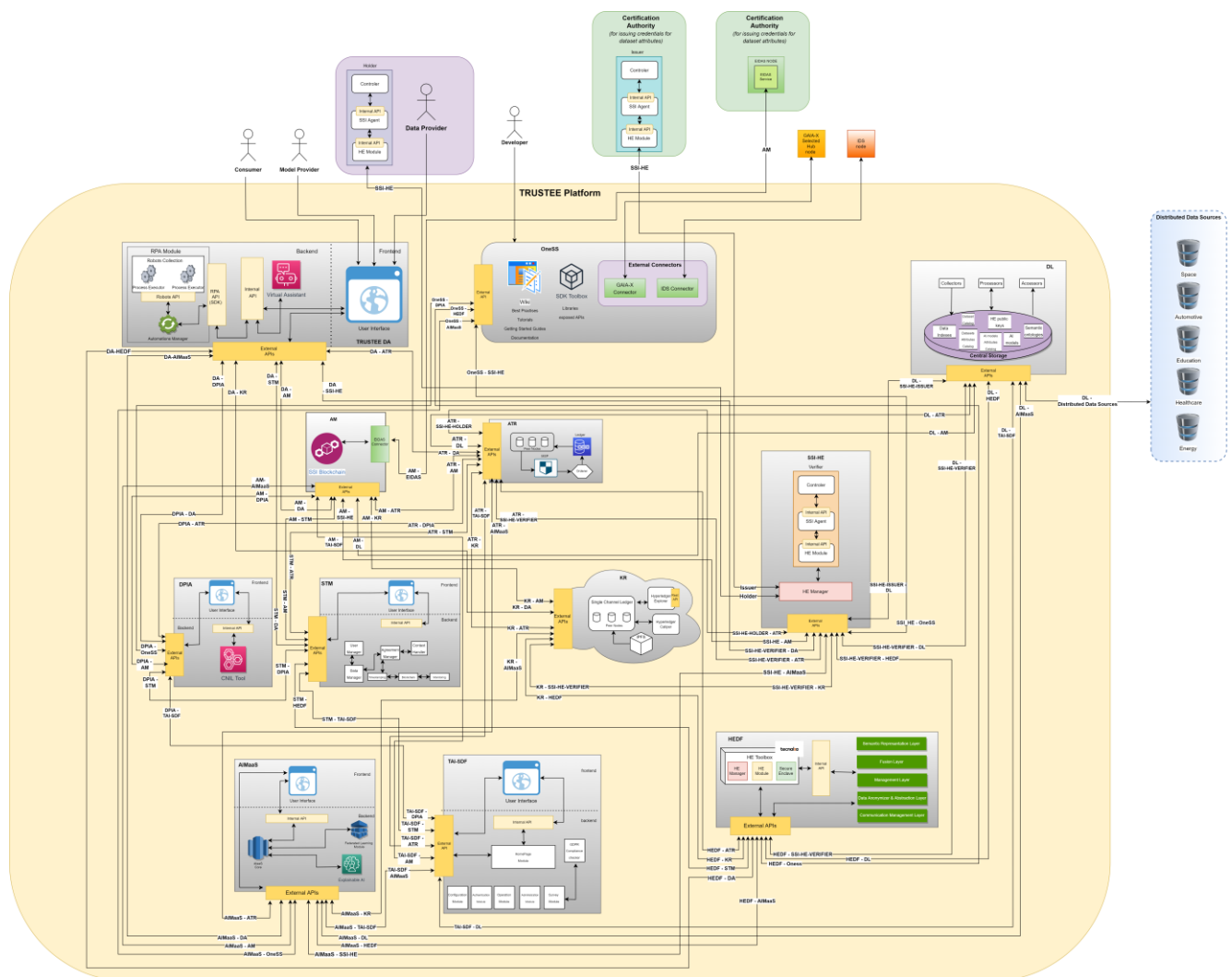


Figure 82: Application Building Blocks overall interaction diagram

BASIC TRUSTEE USE CASE INTERACTION FLOWS

Based on the TRUSTEE Personas developed for each of the Pilot Use Case domains of TRUSTEE, a Basic Use Case Flow has been developed for each Persona aiming to utilize the TRUSTEE Platform, in order to showcase the flow of actions that need to be followed by each Persona to fulfil their primary goal. This flow of actions refers mostly to an initial illustration of the steps that need to be taken for these goals to be realized and not to a

presentation of data flows or to a detailed depiction of the messages that need to be exchanged between the architecture blocks during these flows. Such information will be described in following sections of this deliverable, in the sequence diagrams that have been developed for each ABB and is subject to updates and enhancements in upcoming versions of this deliverable, namely D2.2 and D2.3.

BASIC USE CASE FLOW 1: DATA PROVIDER USE CASE FLOW DIAGRAM

The 1st Basic Use Case Flow refers to a Data Provider Persona that wants to register their data as part of the Distributed Data Sources of TRUSTEE and is showcased in Figure 83 below.

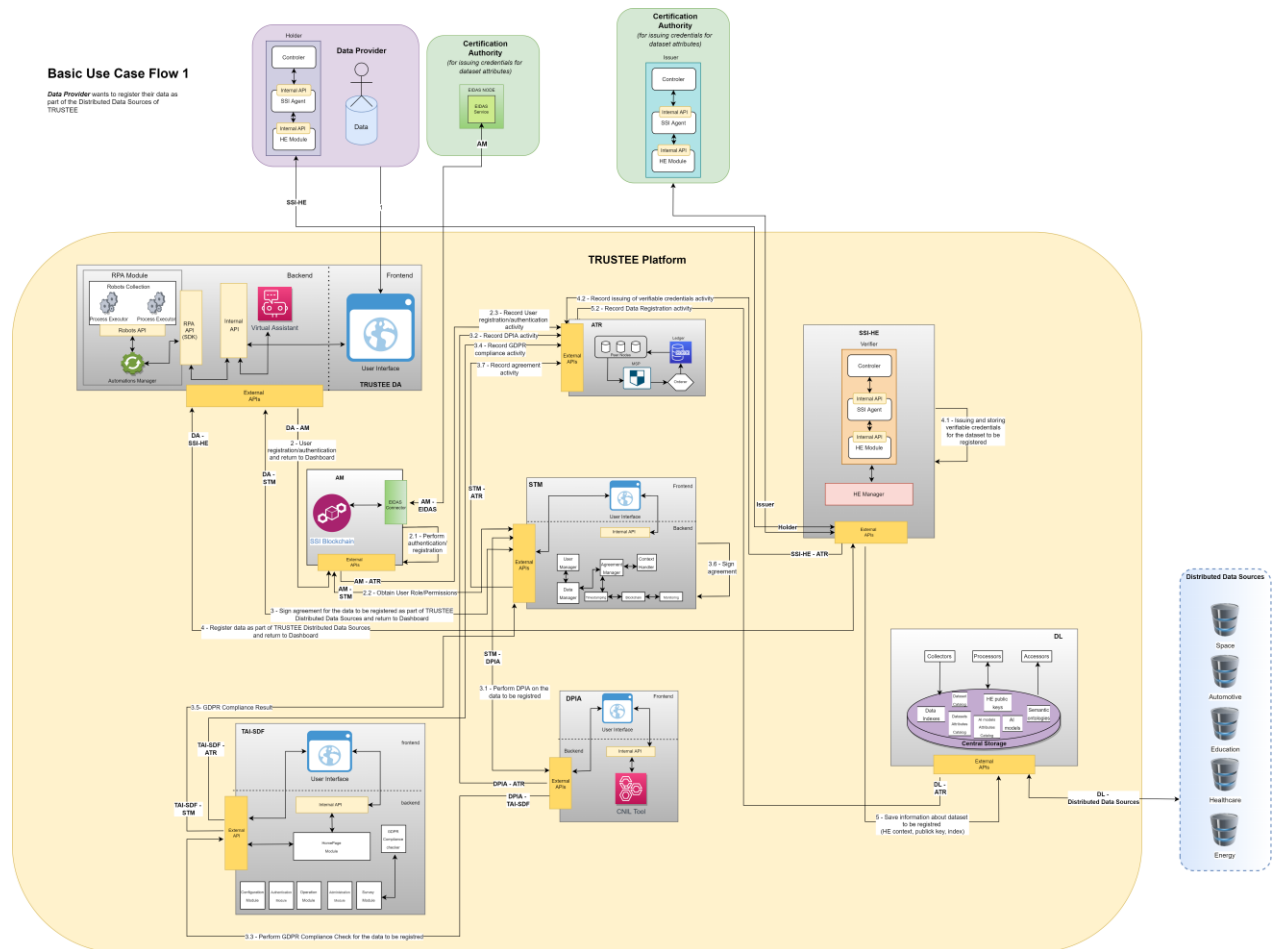


Figure 83: TRUSTEE Data Provider Use Case Flow diagram

BASIC USE CASE FLOW 2: MODEL PROVIDER USE CASE FLOW DIAGRAM

The 2nd Basic Use Case Flow is depicted in Figure 84 and refers to a Model Provider Persona that wants to share their AI/ML models with TRUSTEE.

Basic Use Case Flow 3.1

Data Consumer wants to consume Data that are registered as part of the Distributed Data Sources of TRUSTEE

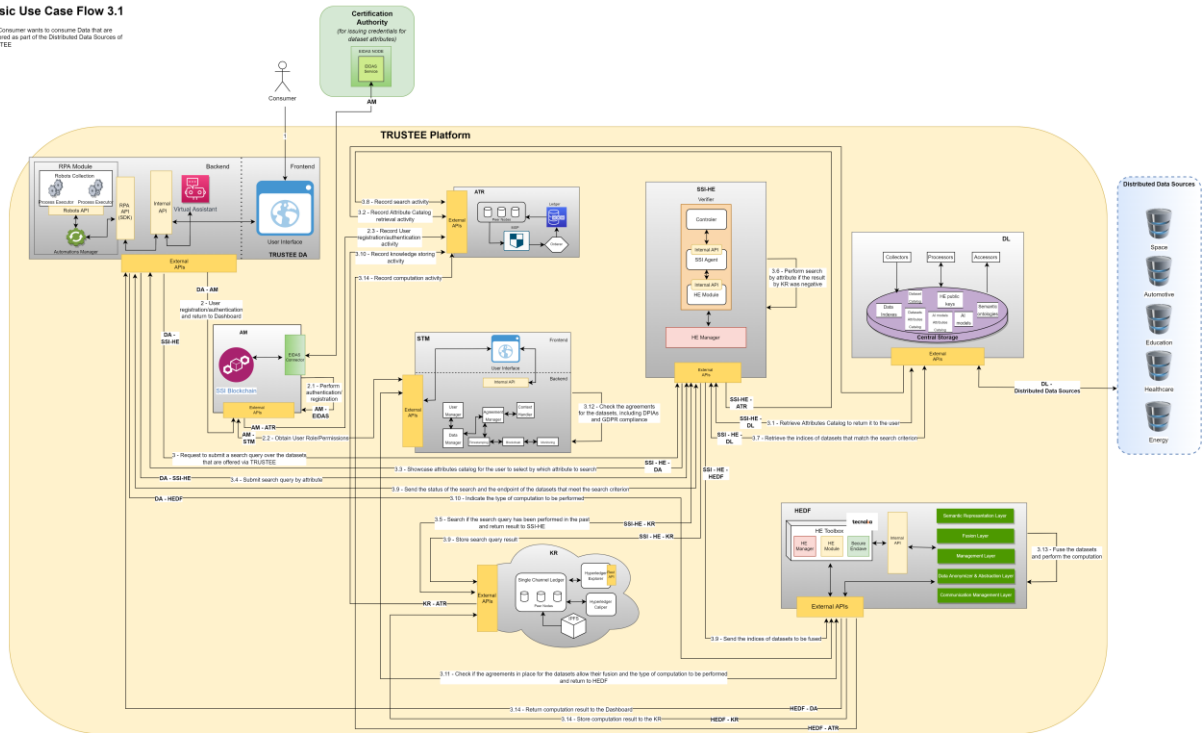


Figure 85: Data Consumer Use Case sub-Flow diagram

Basic Use Case Flow 3.2

Model Consumer wants to consume AML models that are offered by TRUSTEE

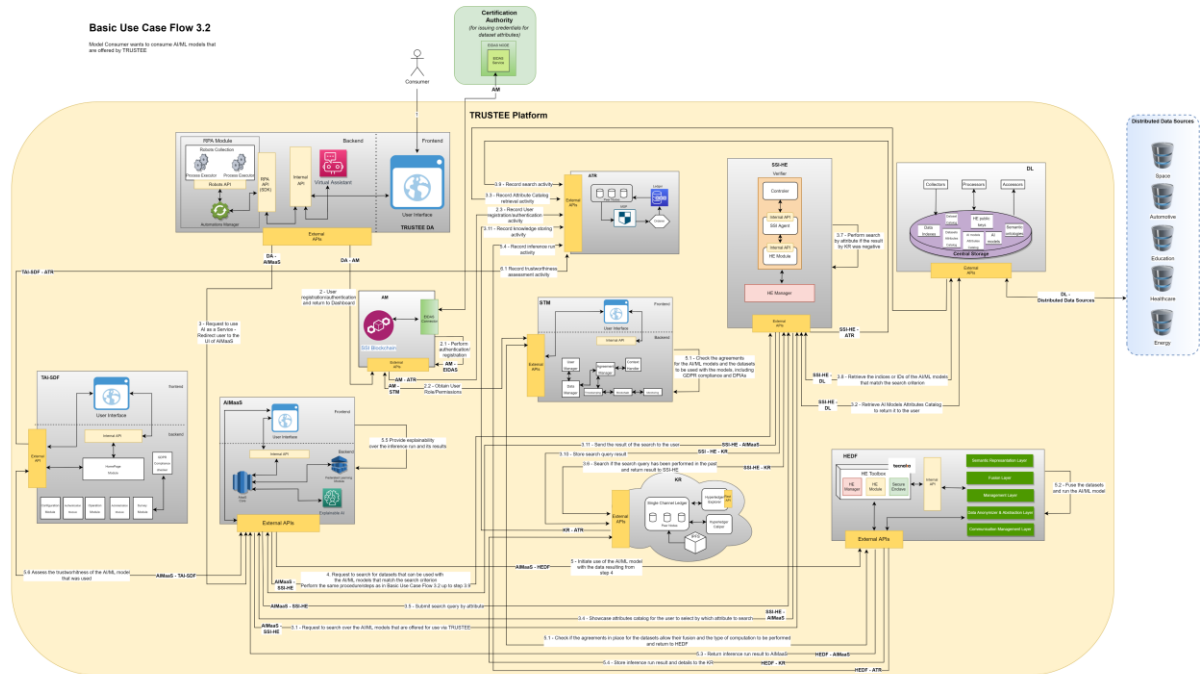


Figure 86: Model Consumer Use Case sub-Flow diagram

BASIC USE CASE FLOW 4: DEVELOPER USE CASE FLOW DIAGRAM

The 4th Basic Use Case Flow refers to a Developer Persona that wants to use functionalities of TRUSTEE exposed via APIs through the One-stop-Shop (OneSS) as part of the incremental strategy of TRUSTEE, or wants to access information offered in the OneSS, such as documentation, tutorials, libraries, and more. This use case flow is presented in Figure 87 and includes five sub-flows depicted in a single diagram, with different color-coding that is explained in the legend at top right corner of the figure, for the shake of avoiding repetition.

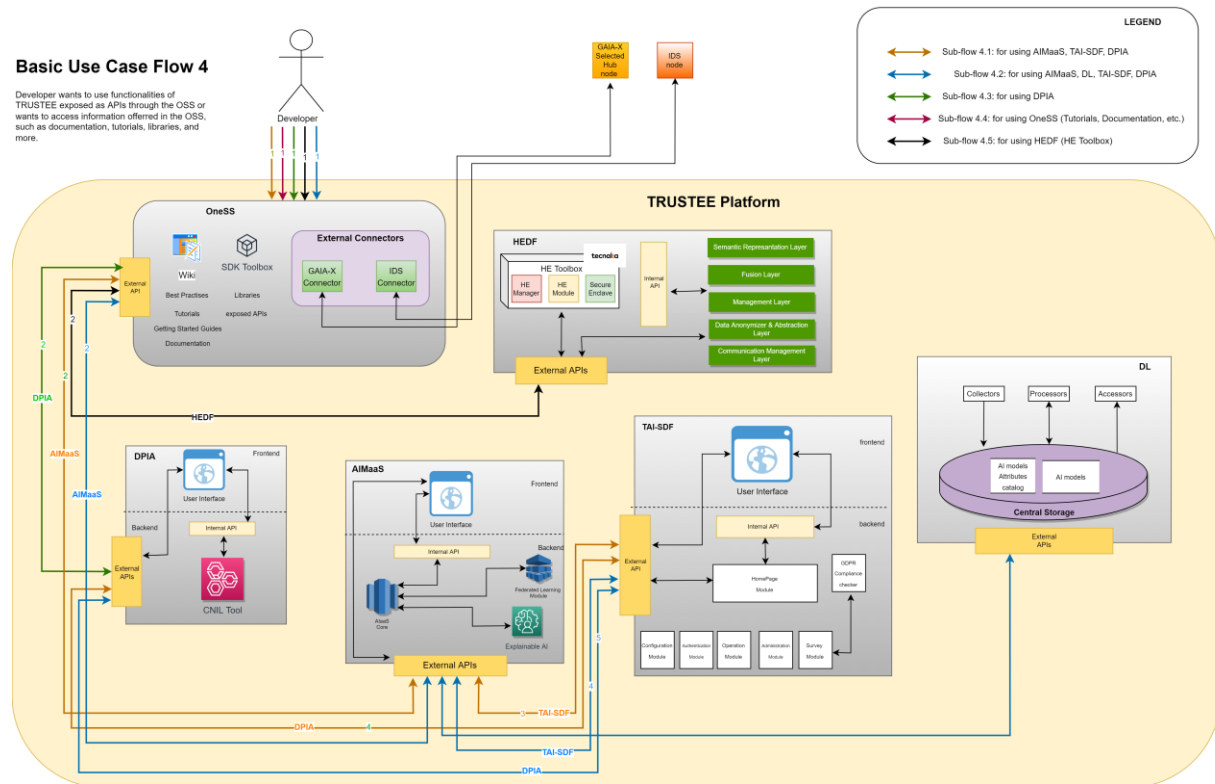


Figure 87: Developer Use Case Flow diagram

USER AUTHENTICATION AND INTERACTION WITH STM AND DPIA SEQUENCE DIAGRAM

A sequence diagram is a type of UML diagram that shows the interactions between objects or components in a system over time. It is typically used to model the flow of messages or events between different components or actors in a system. ArchiMate, however, does not define a sequence diagram explicitly, but it does provide modelling constructs that can be used to represent aspects of a system's behaviour, such as processes, events, and interactions between components. In next figure (Figure 88) we provide an example ABBs sequence diagram, done in Archi tool, modelling the behaviour of TRUSTEE ABBs to illustrate how different TRUSTEE ABBs interact with each other.

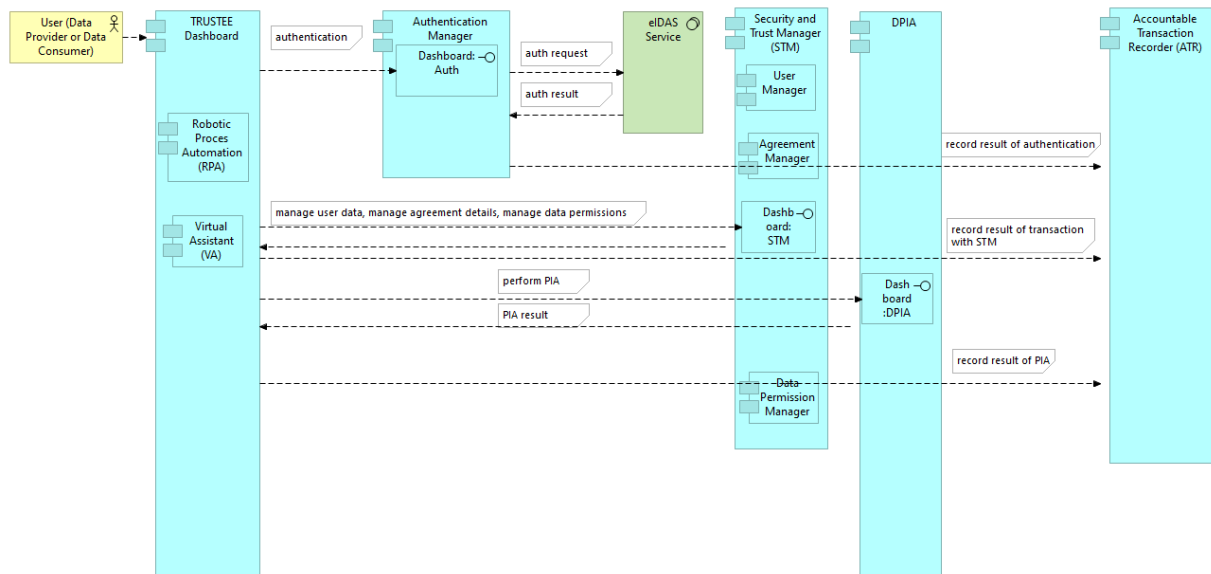


Figure 88: User authentication and interaction with STM and PIA-CNIL components

TRUSTEE ARCHITECTURE BUILDING BLOCKS (ABB)

In this chapter the TRUSTEE ABBs are described. It is expected that some changes may come in the future depending on the newest decisions and alignments. The current version, however, is a final result of multiple iterations, discussions, and alignments between all WP2 partners and TRUSTEE consortium members. For each ABB, we present an ABB overview, ABB application view diagram, and ABB external interfaces specification.

TRUSTEE DASHBOARD (DA) BUILDING BLOCK

ABB overview

The DA ABB provides the main user interface that gives the user the ability to interact with the TRUSTEE Platform. Part of the DA, apart from the the User Interface (UI), are the Virtual Assistant (VA), and the Robotic Process Automation (RPA) modules.

The DA will be a cloud-based, web application that allows the user to fully interact with and utilize the TRUSTEE Platform. The DA will facilitate/visualize for the users of the platform the following processes or will redirect the users to the respective ABB, when necessary:

- Registration and login by communicating with the AM ABB.
- The ability to proceed with performing processes offered by the STM ABB by communicating and redirecting the user to the STM, such as creating agreements over the data they want to register as part of the Distributed Data Sources of TRUSTEE or over the AI/ML models they want to provide to TRUSTEE or creating agreements over data and/or AI/ML models they want to consume or viewing past agreements stemming from the aforementioned actions
- The ability to proceed with performing privacy impact assessments or viewing past assessments by communicating and redirecting the user to the DPIA ABB.
- View their history of activity in the TRUSTEE Platform by communicating with the ATR ABB
- Search operations as well as computational operations over the encrypted data of Data Providers by communicating with the SSI-HE and HEDF ABBs.
- The ability to proceed with the selection and use of ML/AI algorithms by communicating and redirecting the user to the AIMaaS ABB.

- Request to view already stored knowledge stemming from past search queries and data computations by communicating with the KR ABB
- Visualization of the results produced by the end-user's requested computations provided through the communication with the HEDF ABB.
- Navigation assistance and workflow automations through the Virtual Assistant and Robotic Process Automation inner modules.

The VA will be a chat-bot interface with predefined questions and answers that will assist the user to navigate their way into the TRUSTEE Platform. The VA module, through the RPA module, will be capable of providing the user with Automated workflows and Routine procedures to further assist them in using the platform more efficiently and cost-effectively.

The RPA module enables the end-user to interact with the DA in an automated way. Using the VA as the main interface for the RPA processes, the user will be able to request a specific workflow that the VA offers to be performed automatically. Thus, making the interaction with the TRUSTEE platform easier and faster while providing the user with the necessary time to be more productive on other activities. The RPA component comprises:

- **Software Robots** (or Process Executors), which are pieces of software that mimic human actions for interacting with an application interface.
- **Automation Manager** module, which is responsible for the management of the robots.
- **Robot API & SDK**, which is the intercommunication layer that enables the Automation Manager to communicate with the robots and the VA. The Robots SDK provides a way to programmatically create automation processes and deploy robots.

ABB application view diagram

Based on the overview of the DA ABB provided above, an initial version of the DA decomposition diagram is provided below, showcasing the three inner-modules of the block as well as the external API and the initially identified ABBs with which the DA communicates. This diagram is subject to changes, updates, and enhancements as the project progresses, which will be provided in upcoming versions of this deliverable, namely D2.2 and D2.3

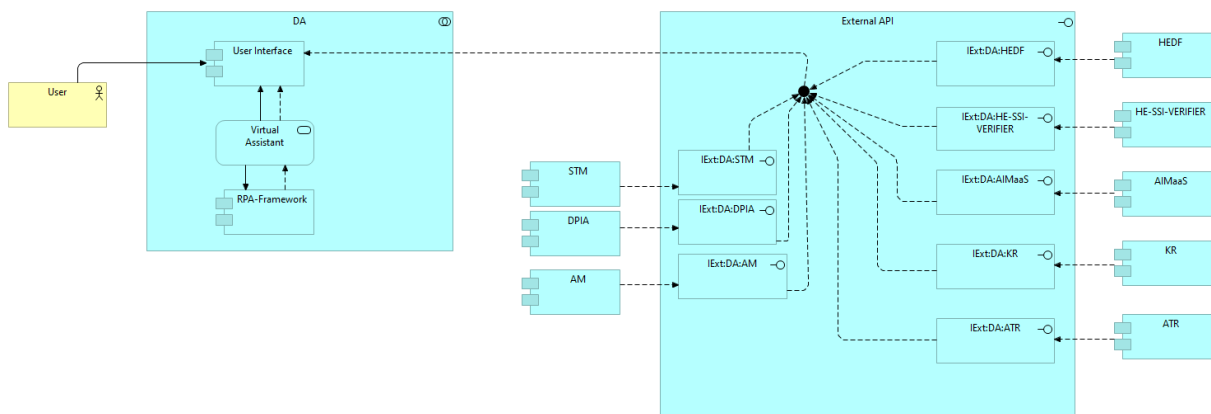


Figure 89: DA decomposition diagram and external interfaces

ABB Sequence Diagrams

For each one of the initially identified external interfaces between the DA and the various ABBs of the TRUSTEE Platform, a sequence diagram has been developed. These sequence diagrams are presented below and will be further updated and discussed in upcoming versions of this deliverable, namely D2.2 and D2.3.

- Facilitating/Visualizing Registration and Login by communicating with the AM ABB can be seen in Figure 90 and Figure 91, respectively.

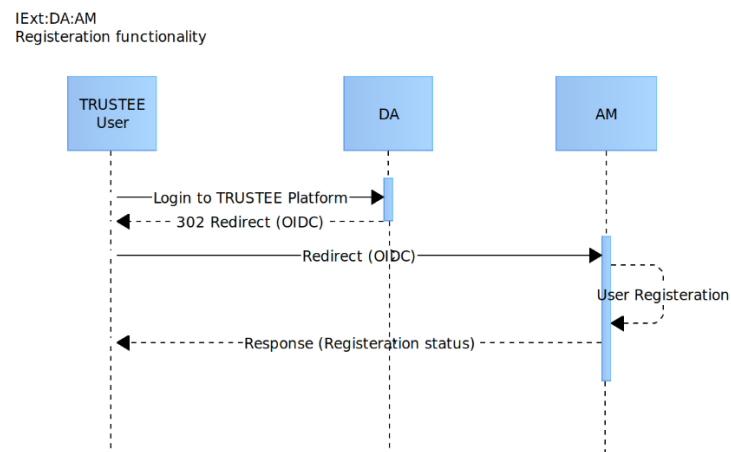


Figure 90: DA-AM register sequence diagram

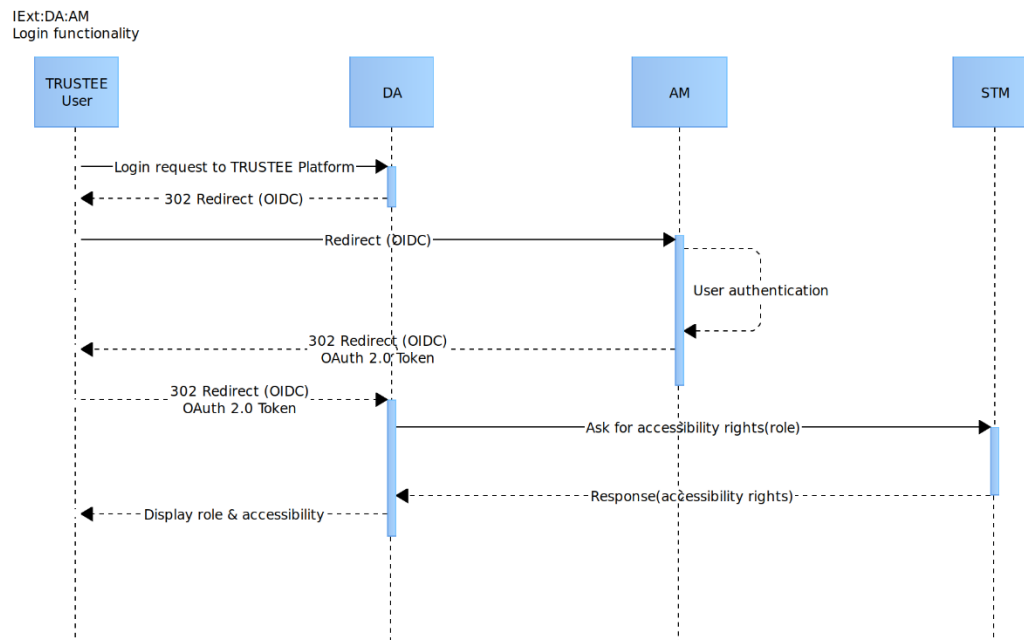


Figure 91: DA-AM login sequence diagram

- Facilitating/Visualizing request and view of history of activity for a user in the TRUSTEE Platform by communicating with ATR ABB is illustrated in Figure 92:

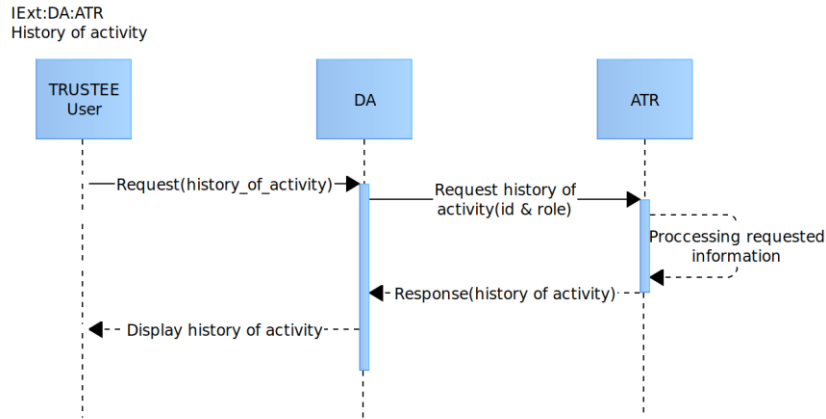


Figure 92: DA request history of activity sequence diagram from ATR

- Facilitating/Visualizing search operations over the encrypted data of the Data Providers by communicating with the SSI-HE ABB is presented in Figure 93.

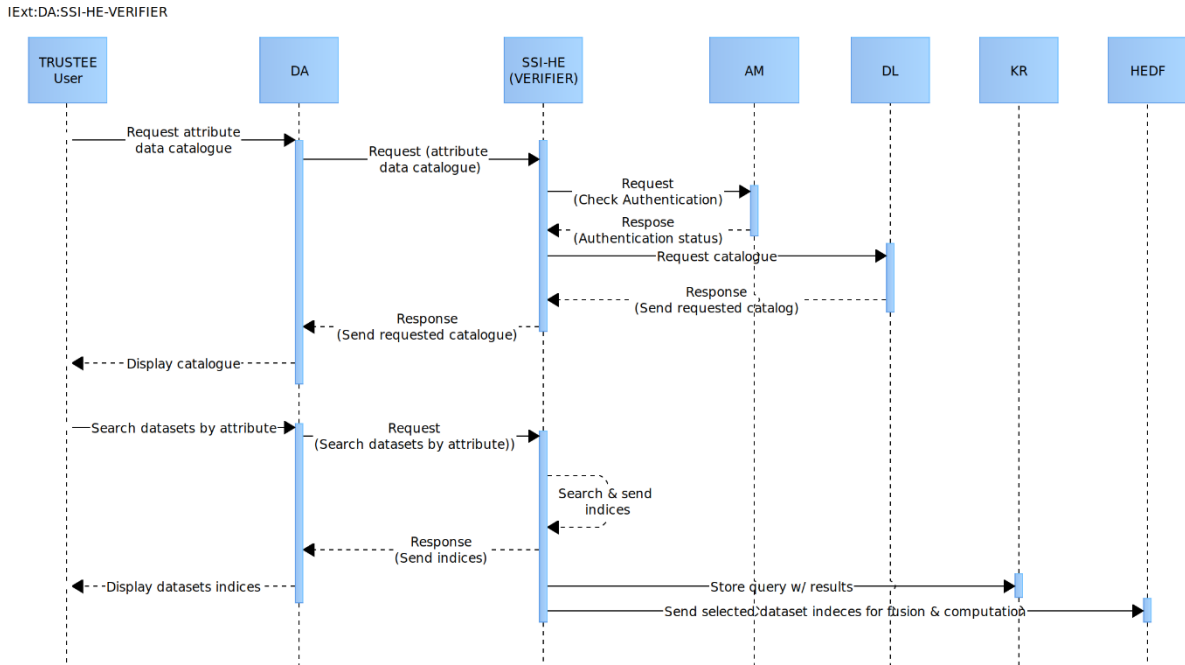


Figure 93: DA HE-SSI-VERIFIER visualizing search operations sequence diagram

- Redirecting the user to the DPIA ABB for performing privacy impact assessments or viewing past assessments is showcased in Figure 94.

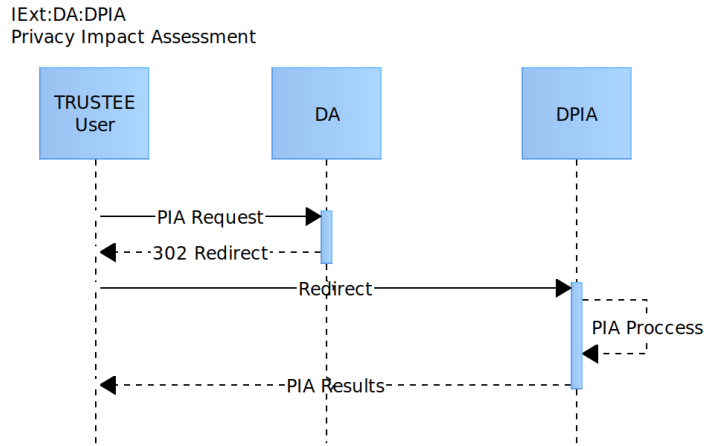


Figure 94: DA-DPIA sequence diagram

- Facilitating/Visualizing the request for viewing stored knowledge stemming from past search queries or results of data computations by communicating with the KR ABB can be seen in Figure 95.

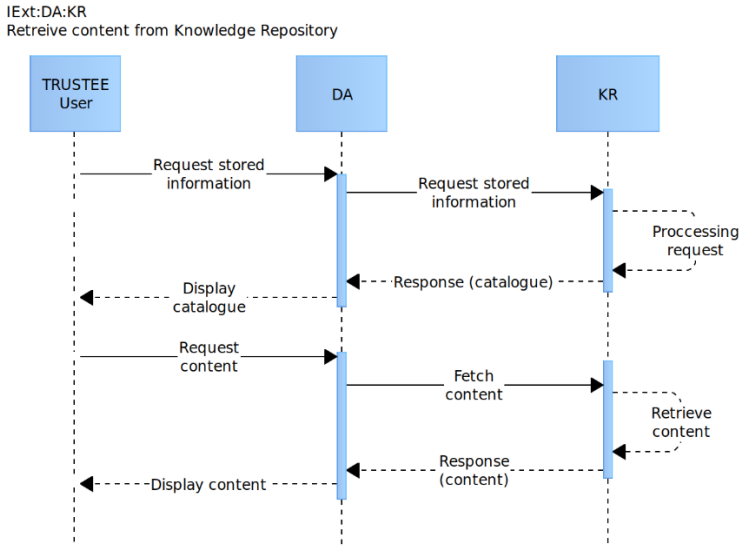


Figure 95: DA retrieve content from KR sequence diagram

- Redirecting the user to the STM ABB for creating or viewing past agreements over data and/or AI/ML models is illustrated in Figure 96.

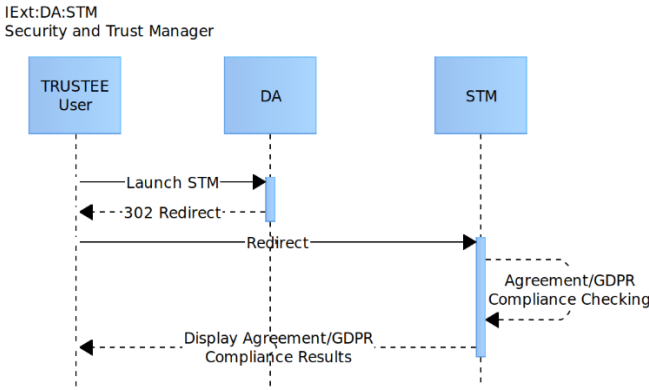


Figure 96: DA-STM sequence diagram

- Redirecting the user to the AIMaaS ABB for performing AI/ML processes is showcased in Figure 97.

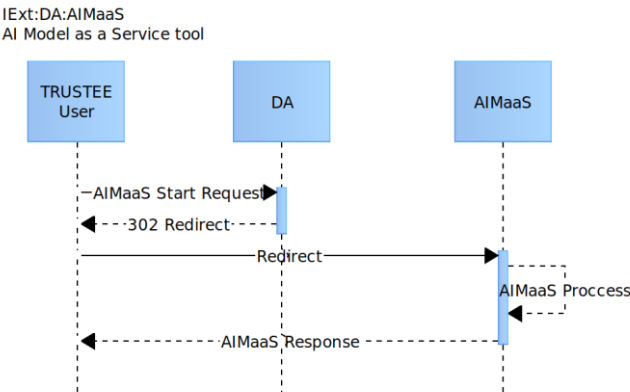


Figure 97: DA-AIMaaS sequence diagram

- Facilitating/Visualizing performance of computations over encrypted data and view of computation results by communicating with the HEDF ABB is presented in Figure 98.

IExt:DA:HEDF

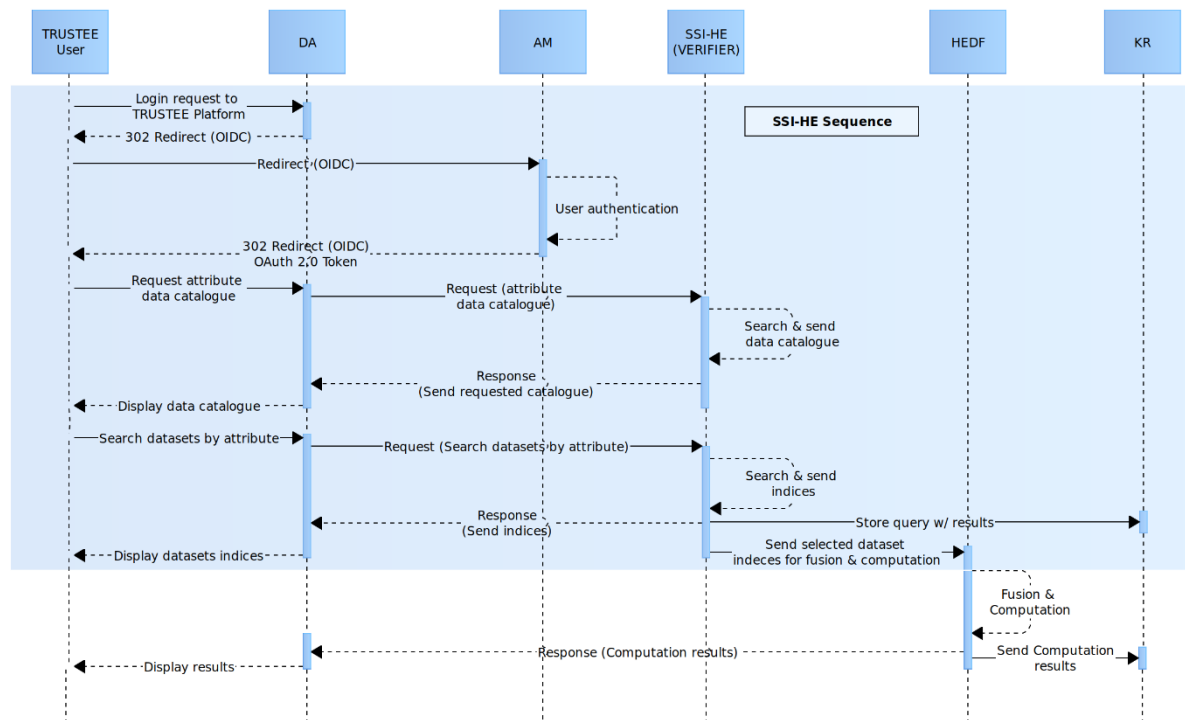


Figure 98: DA-HEDF sequence diagram

ABB External interfaces specification

Based on the overview of the DA component, the initial identification of external communication between the DA and the various ABBs of the TRUSTEE Platform, and the sequence diagrams presented above, below are presented the specifications of the external interfaces of the DA.

Interface	Content	Comment
Interface ID	IExt:DAAuthAgent:AM	
Interface to	Authentication Manager – AM	
Interface overview	The Authentication Agent that will be part of the Dashboard will communicate through this interface with the Authentication Manager ABB when the Dashboard requires to verify that a user is already authenticated or requires to be authenticated before accessing the service, or to authenticate a user	
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	HTTP(S) / OIDC	

Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DA:STM	
Interface to	Security and Trust Manager – STM	
Interface overview	The STM interface will provide users with the necessary agreement and privacy compliance procedures	
Messaging pattern	Navigate/Redirect	
Data exchanged via this interface	The dashboard may need to redirect the user to a particular agreement inside the STM tool, this an identifier for the project/dataset in question should be supplied	
Communication protocol	HTTP(S)	
Message format	URL Query	
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:DA:DPIA	
Interface to	Data Privacy-Impact Assessment - DPIA	
Interface overview	The DPIA interface will provide the user with data privacy impact assessment procedures and with the ability to review their current/past DPIAs	
Messaging pattern	Navigate/Redirect	
Data exchanged via this interface	The dashboard may need to redirect the user to a particular project inside the DPIA tool, this an identifier for the project/dataset in question should be supplied.	
Communication protocol	HTTP(S)	
Message format	URL Query	
Security and integrity	The communication should be done over a Secure HTTP. If the supplied identifiers are leaking information regarding the	

	project/dataset, encryption/anonymisation could be to avoid leak of data in case the user shares the URL to a 3rd party.	

Interface	Content	Comment
Interface ID	IExt:DA:HE-SSI-VERIFIER	
Interface to	SSI-Homomorphic Capable Framework – SSI-HE	
Interface overview	Provide the attribute name and value to be searched and visualize the result of the search	
Messaging pattern	Request/Response	
Data exchanged via this interface	The attribute name and value to be searched as well as the result of the search.	
Communication protocol	HTTPS	
Message format	To Be Defined	Probably JSON
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:DA:ATR	
Interface to	Accountable Transaction Recorder – ATR	
Interface overview	The ATR interface will enable the user to view the history of activity based on their role	
Messaging pattern	Request/Response	
Data exchanged via this interface	Recorded transactional data (user information, events, etc.) & timestamps	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:DA:KR	
Interface to	Knowledge Repository – KR	

Interface overview	The user (consumer) will be able to view knowledge extracted from past queries or use of AI/ML models by them or by other users	
Messaging pattern	Request/Response	
Data exchanged via this interface	Knowledge-related data (e.g., results from past search queries or results of data computations; however, not actual data from the datasets)	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:DA:AIMaaS	
Interface to	AI Models as a Service – AIMaaS	
Interface overview	A TRUSTEE user (consumer) should be able to initiate the AIMaaS tool through the DA	
Messaging pattern	Navigate/Redirect	
Data exchanged via this interface	A ping to start the AIMaaS tool	
Communication protocol	HTTP(S)	
Message format	URL Query	
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:DA:HEDF	
Interface to	Homomorphic Enabled Data Fusion – HEDF	
Interface overview	The Dashboard requests to HEDF the results of the computation that the TRUSTEE User requested	
Messaging pattern	Request/Response	
Data exchanged via this interface	HE data & Data Fusion results	
Communication protocol	HTTP(S)	

Message format	JSON	
Security and integrity	Secure communications (HTTPS)	

AUTHENTICATION MANAGER (AM) BUILDING BLOCK

ABB overview

The TRUSTEE Authentication Manager is the building block tasked to provide authentication related services across the TRUSTEE platform. The AM plays an important role within the TRUSTEE system, providing essential authentication and authorization features for all the TRUSTEE ABBs. By utilizing the AM's external APIs, TRUSTEE ABBs can seamlessly interact with the Authentication Manager, ensuring efficient, user friendly and secure retrieval of user identity information. This integration enables TRUSTEE applications to authenticate and authorize users effectively, enhancing the overall security and trustworthiness of the system.

Moreover, the Authentication Manager serves as a user-facing component, allowing TRUSTEE users to authenticate themselves as TRUSTEE Users using their Self-Sovereign Identity (SSI) wallets. This approach empowers users to have greater control over their identities while providing a secure and user-friendly authentication experience, allowing the user to select what information to share with the systems on each authentication depending on the kind of information requested. Furthermore, by incorporating the eIDAS (Electronic Identification, Authentication, and Trust Services) framework further strengthens the authentication and authorization processes within TRUSTEE. By leveraging eIDAS, TRUSTEE ensures compliance with industry standards and regulations, fostering trust among users and stakeholders. This compliance not only enhances security but also ensures interoperability with other systems and services that adhere to eIDAS standards. Through TRUSTEE users will be able to copy their eIDAS identities to their SSI wallet and use them for authentication and authorisation.

ABB application view diagram

Figure 0.88 shows the Authentication Manager decomposition diagram and displays all the external interfaces along with the other ABBs connected to it. We have to note here that the presented eIDAS nodes, the Stakeholders and the Onboarding Devices are components outside the TRUSTEE ecosystem but connecting with it. Furthermore, please note that in this architecture the Authentication Agents where deployed distributed to take advantage of the distributed nature of the blockchain, thus eliminating the single point of failure. The system will be able to work with a single Authentication Agent.

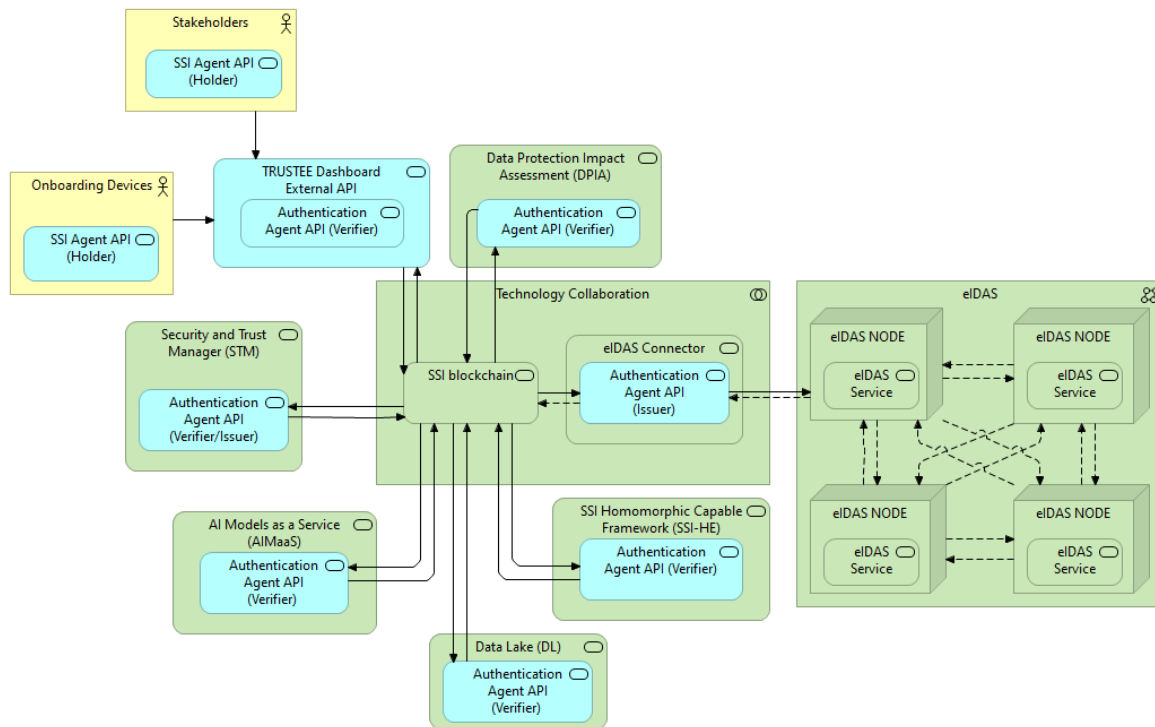


Figure 99: TRUSTEE AM decomposition diagram and external interfaces

ABB Sequence Diagrams

This section presents some sequence diagrams that have been developed for the AM ABB. These diagrams will be further updated and discussed in upcoming versions of this deliverable, namely D2.2 and D2.3.

To start with, the first sequence diagram presented in Figure 100 showcases the onboarding/registration process of a new user in the TRUSTEE ecosystem through the use of the AM ABB and by utilising eIDAS. The process starts as a normal login process, but at the point where the user is asked for proof of their identity, a registration option is also provided to allow the onboarding of new users. In this registration process presented, the user's identity is retrieved through the eIDAS ecosystems, and the appropriate attributes are issued as SSI verifiable credentials by the AM to the user for them to store on their wallet.

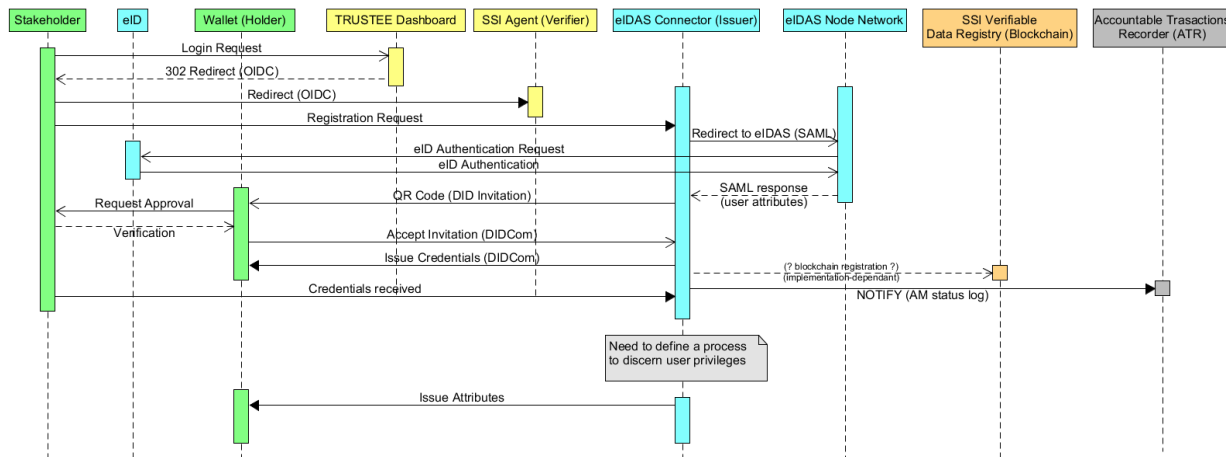


Figure 100: TRUSTEE AM onboarding (registering) new user on the TRUSTEE system.

The next sequence diagram presented in 90 illustrates the login process for a user already registered in the TRUSTEE Platform. The Login request is initiated by an ABB (in this case the TRUSTEE Dashboard) and through OpenID Connect the AM is able to verify the user's SSI identity and deliver it back to the ABB.

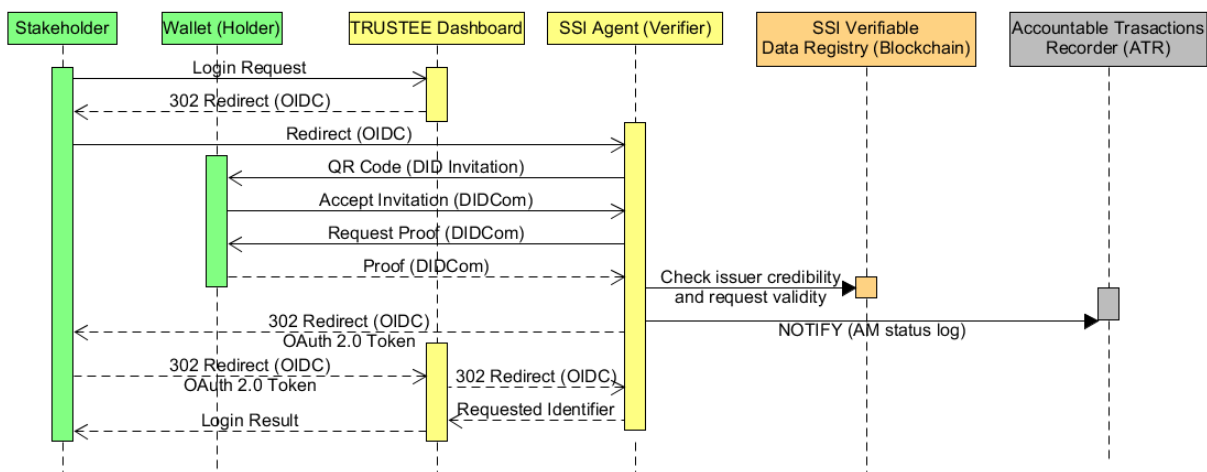


Figure 101: TRUSTEE AM authenticating (login) user on TRUSTEE's Dashboard.

ABB External interfaces specification

Interface	Content	Comment
Interface ID	IExt:AM:ABBAuthAgent	

Interface to	TRUSTEE Dashboard - DA Security and Trust Manager -STM Data Protection Impact Assessment - DPIA AI Models as a Service – AIMaaS SSI Homomorphic Capable Framework (SSI-HE) Data Lake – DL Trustworthiness AI Support Design Framework - TAI-SDF	<i>Many of the TRUSTEE ABBs will need to use this interface to authenticate users.</i>
Interface overview	The interface is associated to an Authentication Agent that as a software will be part of the ABBs which will require to verify that a user is already authenticated or requires to be authenticated before accessing the service, or to authenticate a user.	
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	HTTP(S) / OIDC	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:AM:ATR	
Interface to	Accountable Transactions Recorder – ATR	
Interface overview	This interface will be used for logging AM transactions to the ATR ABB	
Messaging pattern	<i>Request/response, publish-subscribe, notification</i>	
Data exchanged via this interface	<i>Transaction data</i>	
Communication protocol	<i>e.g., gRPC, HTTP(S), REST API, FABRIC API</i>	
Message format	<i>JSON, XML, CSV, protobuf, CBOR, BSON.</i>	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:AM:Agent2Agent	
Interface to	TRUSTEE AM Agent External SSI Agent	
Interface overview	Communication between 2 TRUSTEE SSI Agents. Depending on the agent and the operation to be performed this interface could facilitate various activities (e.g. messaging, presentation of attributes, issue of verifiable credentials)	
Messaging pattern	Request/response	
Data exchanged via this interface	DID data	
Communication protocol	DIDComm/HTTP(S)	Authentication Agents use the DIDComm protocol to communicate between each other. DIDComm can be transmitted over any channel (e.g. HTTPS)
Message format	DIDComm	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:AM:SSIEIDAS	
Interface to	SSI eIDAS Trusted Issuer	
Interface overview	The service responsible for linking SSI identities with the eIDAS will have to communicate with the eIDAS network. For this, a connection will have to be made towards the appropriate eIDAS node.	
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	SAML	
Message format	eIDAS SAML Message Format	
Security and integrity		

SECURITY AND TRUST MANAGER (STM) BUILDING BLOCK

ABB overview

The STM will enable the TRUSTEE stakeholder to have total control over their data providing. It is a usable and workable trust, stakeholder control, and GDPR Compliance-based explicit agreement module that provides the procedure for collecting and use of the data and (AI models), policies and procedures to protect that data and/or AI models.

STM is a web-based application that allows the user to fully interact with it for managing the agreement via an agreement manager. The agreement manager takes into consideration the creation and generation of an agreement from the Data and/or Model provider. An agreement is requested and signed by a Consumer. The main functionalities that the STM will provide to the users of the platform:

- Create a user: Implement a user creation (registration) to manage users within the STM and associate them to agreements. This functionality will be communicating with the DA and AM ABBs.
- Authentication: Enable users that already exist within the STM to maintain and manage their agreements.
- Create a dataset agreement (as Data Provider): Allow to provide a user-friendly web-based UI for creating data agreement. It also provides a backend with a NoSQL database (GraphDB) for storing data agreement.
- Data search: Provide a user-friendly web-based UI to facilitate the data search based on the meta-data.
- View Dataset: Display all of the information related to the user's selected dataset
- Accepting/Signing the agreement: Provide a user-friendly web-based UI and backend to accept/sign the agreement.
- View Agreements: Provide a user-friendly web-based UI to display the users currently accepted agreements.
- View Profiles: Enable a user to view their own profile information and/or the profile of a Data Provider.

ABB application view diagram

The following diagram shows the general principal of the STM. Given a data subject sharing their data with corresponding consent to a Data Provider (out of the scope of TRUSTEE). STM is a component facilitating the data sharing and agreement life-cycles between a Data Provider and a Consumer of the TRUSTEE Platform.

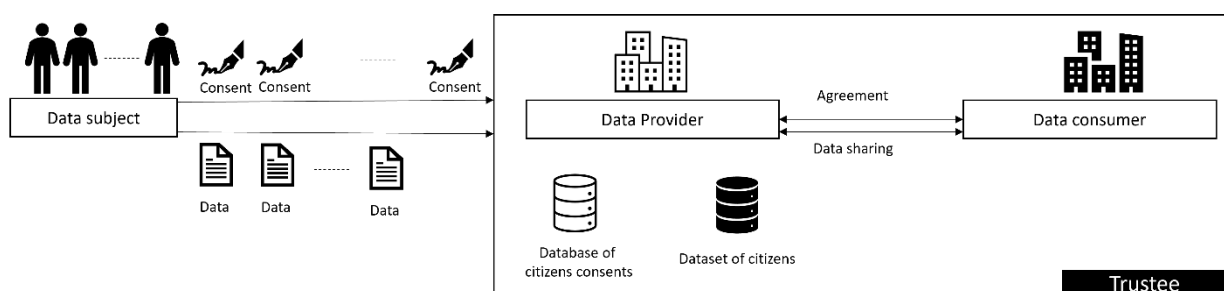


Figure 102: STM context

Figure 102 below shows an initial version of the data flows regarding STM. It is composed of the agreement manager responsible for multiple operations and functionalities, such as the data and agreement life cycles like creating, editing, accepting the agreements. The agreement manager is connected to the ATR, HEDF, AM, and

DPIA ABBs. More details of the STM will be included in D4.1 and further updates will be provided in future versions of this deliverable, namely D.2.2 and D2.3.

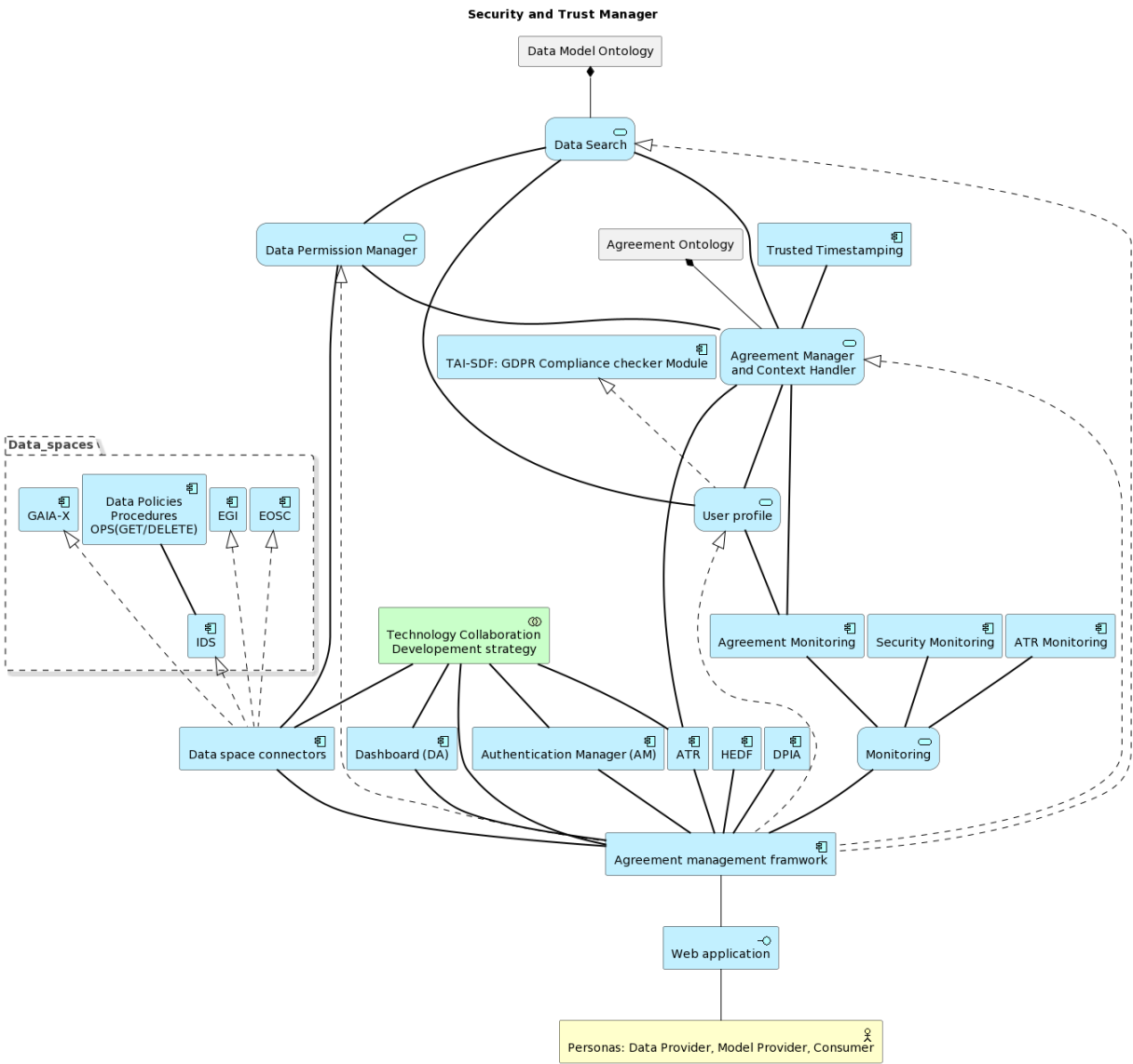


Figure 103: STM data flows

ABB Sequence Diagrams

The sequence diagrams have been developed for the STM ABB highlighting the interactions between Data Providers and Consumers when using the STM. It also shows the interactions between the STM and other ABBs such as AM, ATR, HEDF, DPIA, and TAI-SDF.

The following figure shows the interactions between the Data Provider and the STM when a Data Provider wants to create an agreement for a dataset. This agreement will be accepted or signed by a Consumer for a data and/or AI model use.

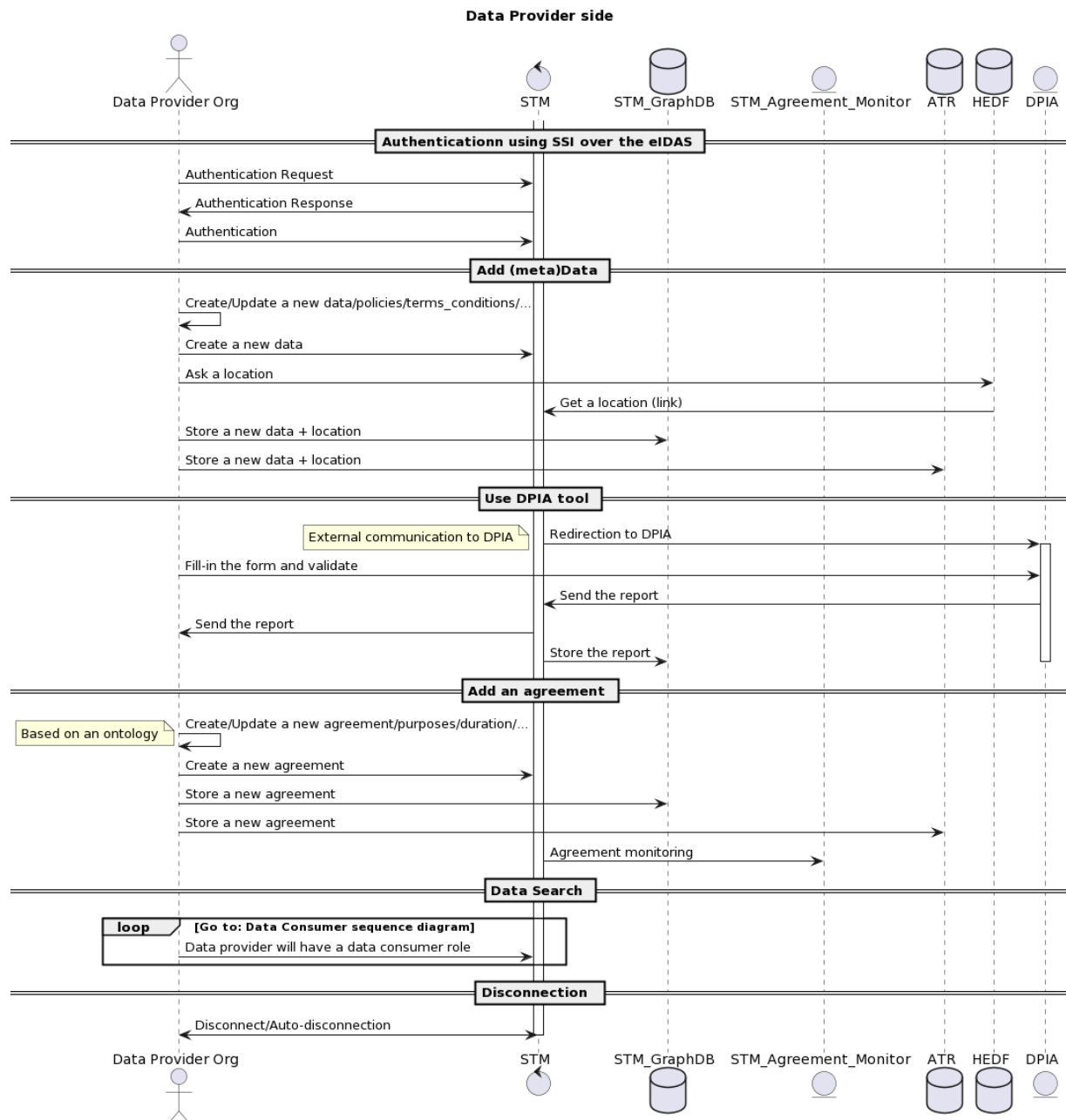


Figure 104: STM and Data Provider interactions while creating an agreement for a dataset

The following figure shows the interactions between the Consumer and the STM when a Consumer wants to use data and/or AI models. First, the Consumer searches the data based on a meta data or keywords. Second, the STM shows the existing data corresponding to the user's request. Third, the Consumer accepts the agreement of the dataset and/or AI model in order to use it.

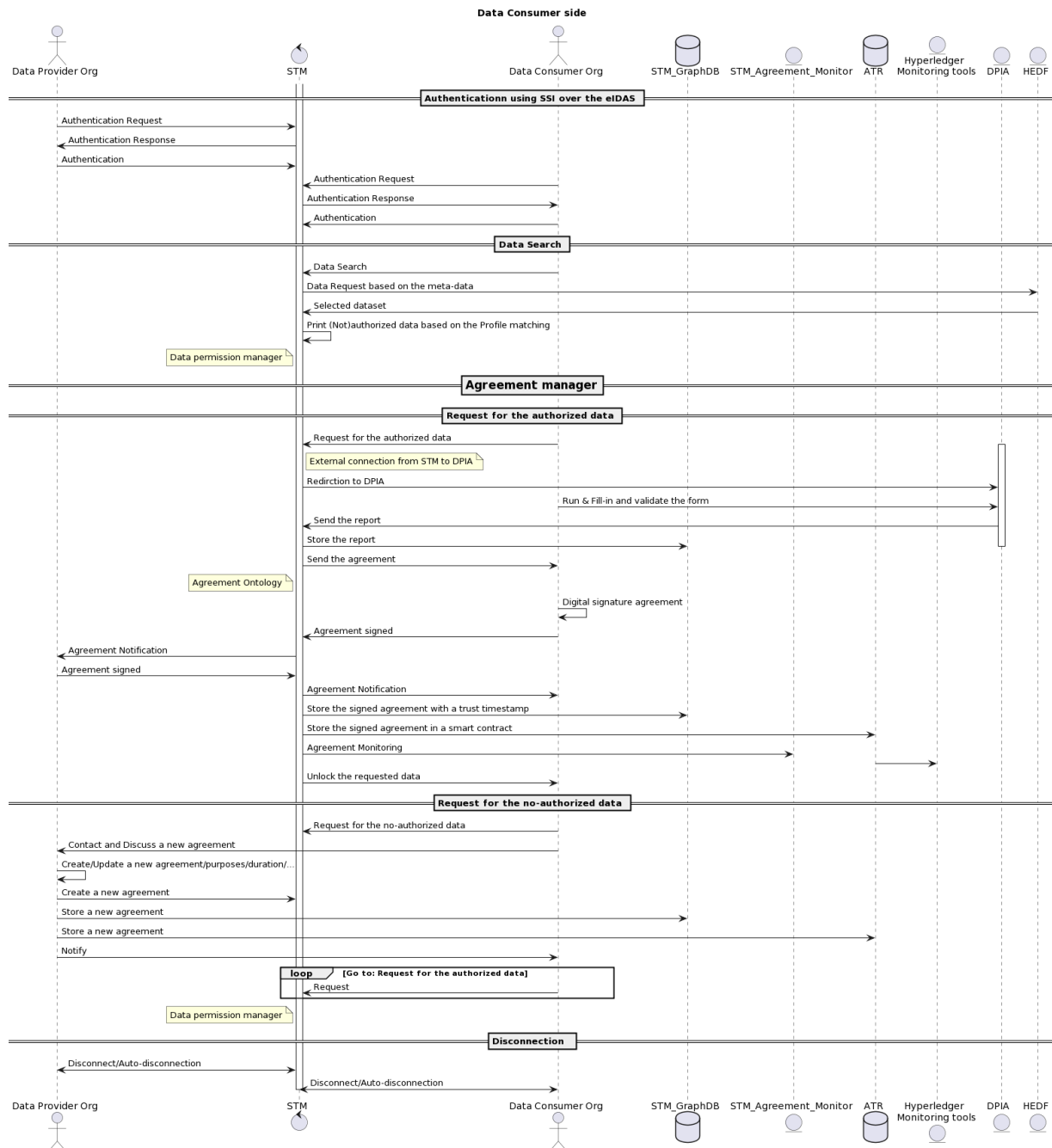


Figure 105: STM and Consumer interactions while using data and/or AI models

The following figure shows the the agreement validity checker of the STM. The goal of the agreement validity checker is to check if the agreement is valid based on the GDPR requirements such as the duration of the agreement, purposes, actions, and location.

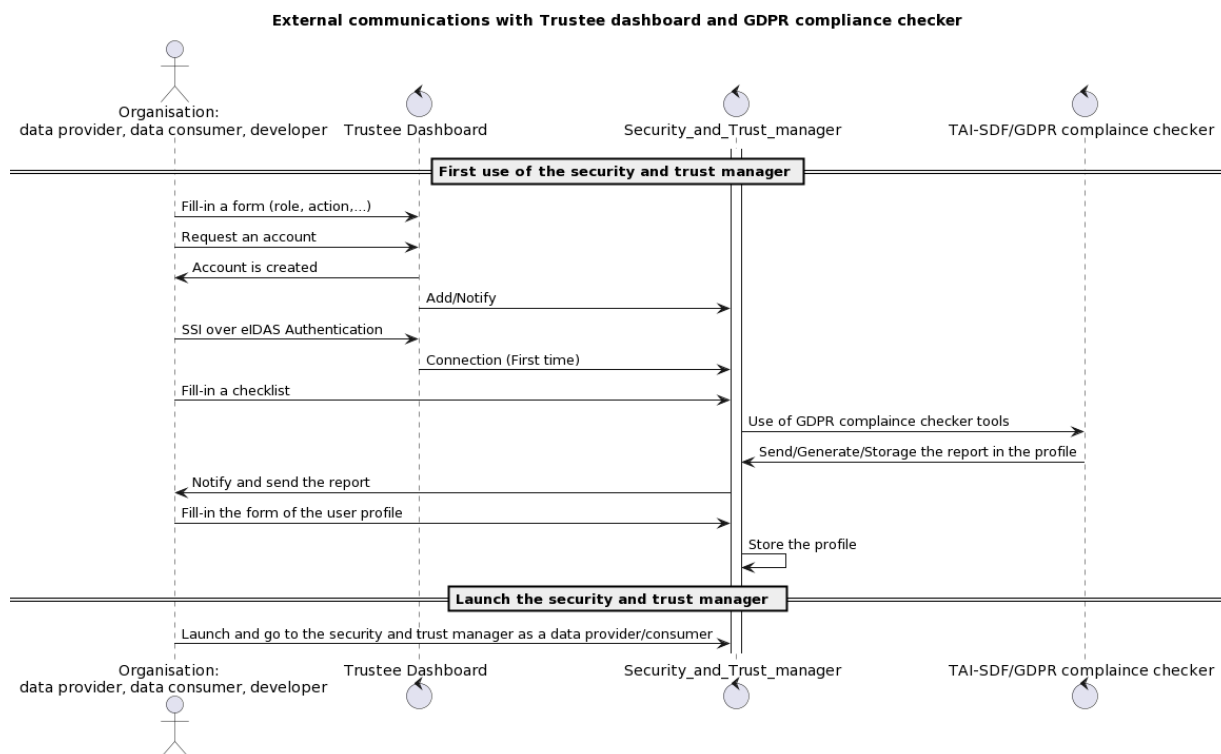


Figure 107: STM sequence diagram, communication with DA and compliance checker

ABB External interfaces specification

Interface	Content	Comment
Interface ID	IExt:STMAuthAgent:AM	Used for authentication
Interface to	Authentication Manager – AM	
Interface overview	The goal is to use a secured authentication interface using SSI over the eIDAS for connecting to the agreement management platform	
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	HTTP(S) / OIDC	
Message format	JSON	

Security and integrity		

Interface	Content	Comment
Interface ID	IExt:STM:ATR	Connection between the Security and Trust manager and the blockchain(ATR)
Interface to	Accountable Transactions Recorder – ATR	
Interface overview	The goal is to use TRUSTEE Accountable Transactions Recorder (ATR) for recording the agreement between the data provider and consumer	
Messaging pattern	publish-subscribe	
Data exchanged via this interface	A record: <agreementID, data providerID, data consumerID, timestamps, public key, datasetID>	
Communication protocol	Free to secure the communications	
Message format	JSON	
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	The data transmission should be secured.

Interface	Content	Comment
Interface ID	IExt:STM:HEDF	
Interface to	Homomorphic Enblaed Data Fusion – HEDF	
Interface overview	STM checks for agreements on datasets and/or agreements on models, privacy policies and DPIA regarding a specific computation that is going to be performed by the HEDF. The check will be requested by HEDF.	
Messaging pattern	Request/response	
Data exchanged via this interface	Check Status/Outcomes	
Communication protocol	Free to secure the communications	
Message format	JSON	
Security and integrity	HEDF, upon a data consumer's request, checks from STM agreements on datasets, agreements	

	on models, privacy policies and DPIA regarding a specific computation.	

Interface	Content	Comment
Interface ID	IExt:STM:DA	
Interface to	Dashboard – DA	
Interface overview	The goal is to use the Trustee dashboard for launching the security and trust manager.	
Messaging pattern	Request/response	
Data exchanged via this interface	A ping to launch the security and trust manager	
Communication protocol	HTTP(S)	
Message format	URL Query	
Security and integrity	Over internet or local infrastructure (private cloud), or Trustee SDK, or Trustee platform	The exchanged information should be secured.

Interface	Content	Comment
Interface ID	IExt:STM:TAI-SDF	GDPR compliance checker will be used
Interface to	Trustworthy AI Support Design Framework – TAI-SDF	
Interface overview	The goal is to use the GDPR compliance checker for assessments	
Messaging pattern	Request/response	
Data exchanged via this interface	JSON/PDF	
Communication protocol	HTTPS, or others	
Message format	Not yet defined	
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	

Interface	Content	Comment
Interface ID	IExt:STM:DPIA	

Interface to	Data Privacy Impact Assessment – DPIA	
Interface overview	The goal is to use the DPIA for each data request in order to check if the agreement matches the privacy impact assessment.	
Messaging pattern	Request/response	
Data exchanged via this interface	JSON	
Communication protocol	HTTPS, or others	
Message format	Not yet defined	
Security and integrity	Trustee platform	

DATA PROTECTION IMPACT ASSESSMENT (DPIA) BUILDING BLOCK

ABB overview

The Data Protection Impact Assessment (DPIA) ABB is offering DPIA services, allowing users to insert information about the data, the intended processing and by considering the external actors and threats, to create a report. Normally such a service doesn't interact with other components of a system as it acts independently following the processes defined by each organisation. In TRUSTEE the DPIA ABB will be enhanced and allow interaction with other ABBs. DPIA will communicate with the ATR and the STM sharing probably the status of each assessment and provide access to the information inside.

The DPIA ABB will include:

- Legal and ethical knowledge base for ensuring the lawfulness of processing and protection of data subjects' rights.
- Contextual knowledge base, available along all the steps of the PIA, adapting the content displayed.
- User-friendly interface allowing for a simple management of the PIA.

The communication between the DPIA ABB and the SSI-HE could be characterized as "Future enhancements". The only reason PIA would communicate with the HE framework would be the "creation" of a new dataset by the HE framework. In that case the HE framework would take the role of a dataset owner, that needs PIA to report that this new dataset is compliant with GDPR.

ABB application view diagram

The following Figure 108 presents the DPIA ABB's decomposition diagram presenting its main internal components and interconnecting it with other ABBs through the external interfaces.

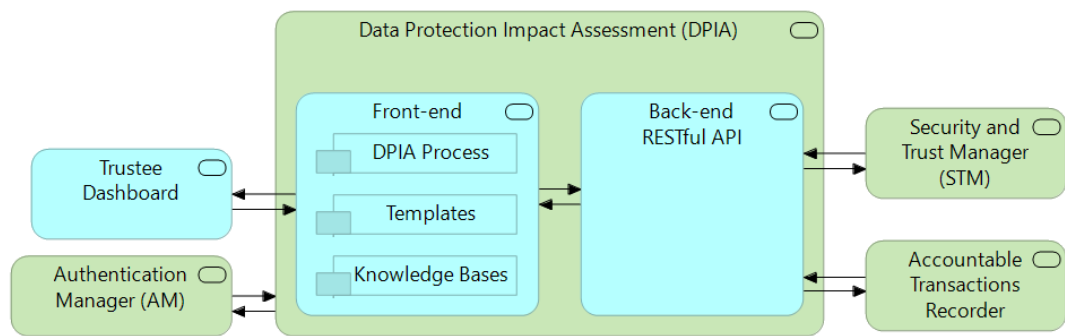


Figure 108: TRUSTEE DPIA decomposition diagram and external interfaces.

Finally, the Figure 109 below presents in more details the internal components of the DPIA ABB and the interfaces with external ABBs.

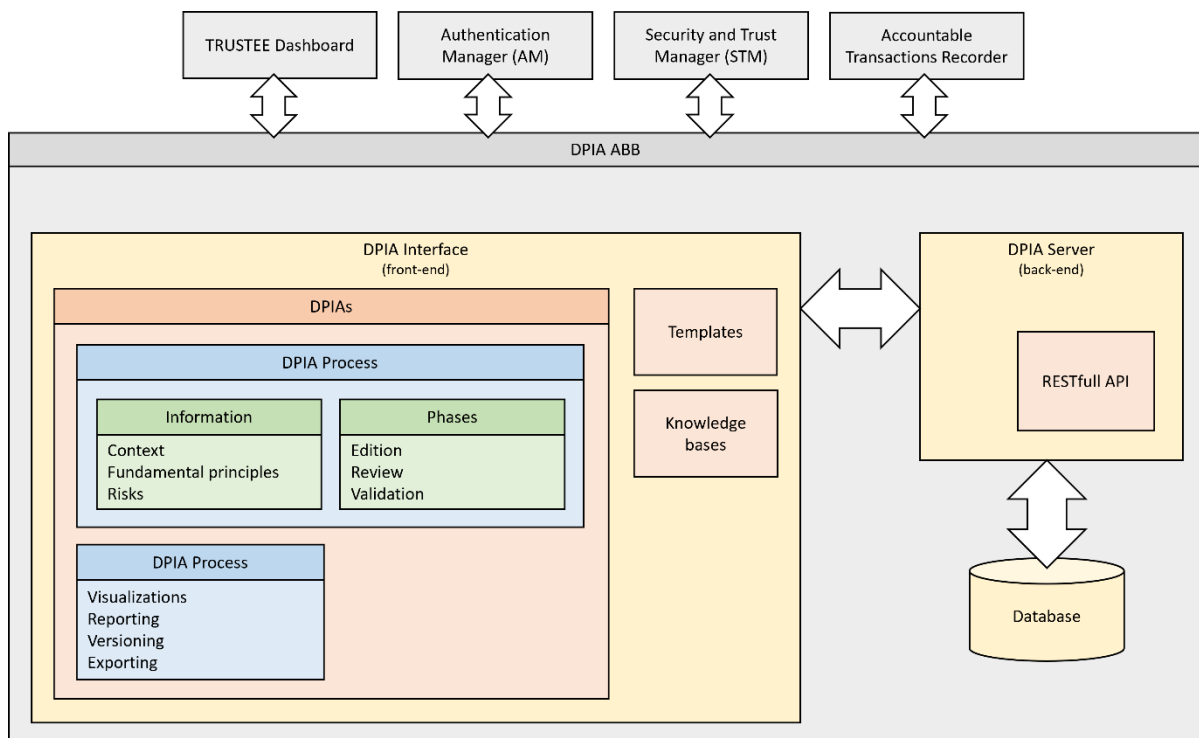


Figure 109: DPIA context and breakdown

ABB Sequence Diagrams

In this section we present the main sequence diagram of using the DPIA ABB. Figure 110 shows the process a TRUSTEE user has to follow to access the DPIA ABB, authenticate, start a new DPIA, work on it filling in the information needed, review the DPIA and finally approve or reject it. Furthermore, the flow showcases how other ABBs can access information regarding the outcome of a DPIA to be used for their services.

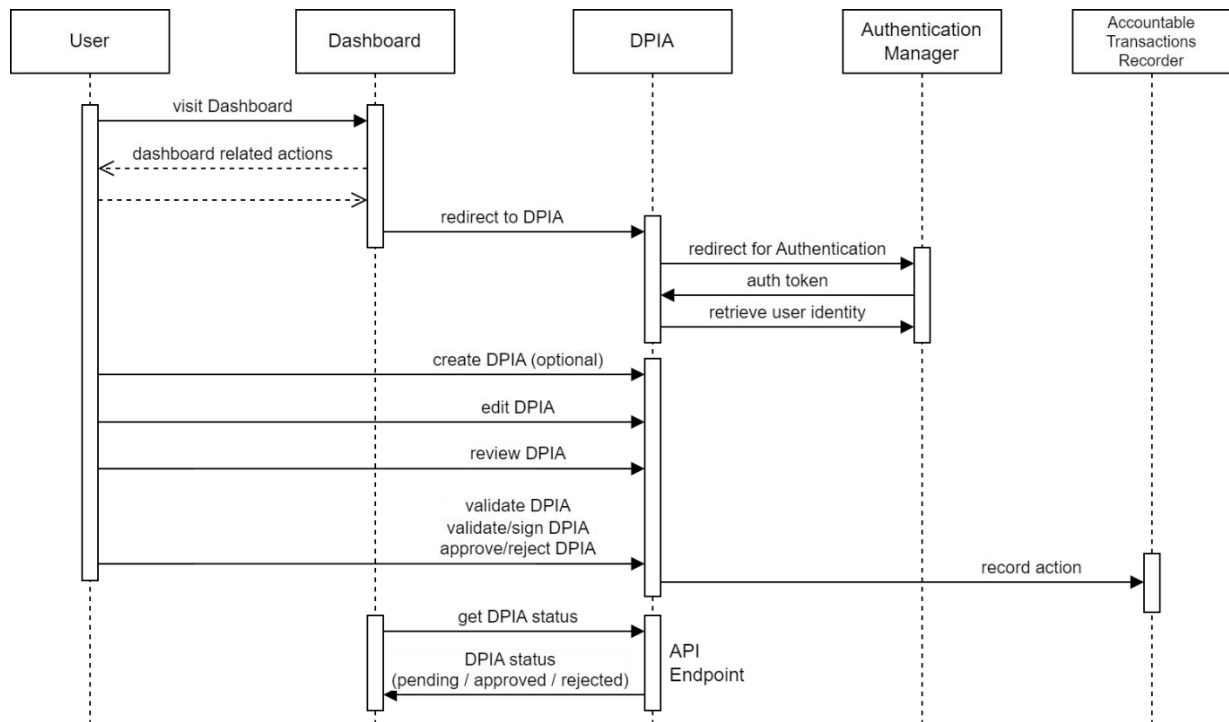


Figure 110: TRUSTEE DPIA user creating, editing, reviewing, validating/signing/rejecting assessment and ABB recovering DPIA status.

ABB External interfaces specification

Interface	Content	Comment
Interface ID	IExt:DPIA:DA	
Interface to	Dashboard – DA	
Interface overview	A TRUSTEE user should be able to access the DPIA tool through the DA.	
Messaging pattern	Navigate/Redirect	
Data exchanged via this interface	The dashboard may need to redirect the user to a particular project inside the DPIA tool, this an identifier for the project/dataset in question should be supplied.	
Communication protocol	HTTP(S)	

Message format	URL Query	
Security and integrity	The communication should be done over a Secure HTTP. If the supplied identifiers are leaking information regarding the project/dataset, encryption/anonymisation could be to avoid leak of data in case the user shares the URL to a 3rd party.	

Interface	Content	Comment
Interface ID	IExt:DPIA:ATR	
Interface to	Accountable Transactions Recorder – ATR	
Interface overview	Interaction with the DPIA tool that result in change of the DPIA results or access to sensitive data should be logged.	<i>To be investigated further what should be recorded, e.g., Project result generation, by which user</i>
Messaging pattern	Request/response	
Data exchanged via this interface	An identifier of the user/entity that interacted with the tool. An identifier of the DPIA accessed/modified. The type of the action performed (e.g., access, modification). A timestamp of the event.	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	Both the DPIA tool and the Accountable Transactions Recorder should be authenticated during this interaction. The message integrity (to avoid tampering) and authenticity (that was generated by the DPIA tool) should be ensured.	

Interface	Content	Comment
Interface ID	IExt:DPIAAuthAgent:AM	
Interface to	Authentication Manager – AM	

Interface overview	The DPIA tool will have to authenticate the user before giving him access to the project/data.	<i>To be further investigated regarding whether to authenticate the user and check if he has access to the project or verify that the user has permission to access the data (verifiable credentials)</i>
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	HTTP(S) / OIDC	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DPIA:STM	
Interface to	Security and Trust Manager – STM	
Interface overview	The DPIA tool will have communicate the results of the DPIA with the Security and Trust Manager.	
Messaging pattern	Request/response	
Data exchanged via this interface	The project/dataset identifier. The status or the DPIA (e.g., created, pending, completed, updated). The results of the DPIA.	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	Both the DPIA tool and the Accountable Transactions Recorder should be authenticated during this interaction. The message integrity (to avoid tampering) and authenticity (that was generated by the DPIA tool) should be ensured.	

ACCOUNTABLE TRANSACTION RECORDER (ATR) BUILDING BLOCK

ABB overview

- The ATR will be implemented utilizing a **permissioned** blockchain technology; thus, only the authorized entities/nodes will be able to join the blockchain network.
- Hyperledger Fabric, a permissioned blockchain framework, is a privacy-respecting technology which mitigates several crucial privacy and security threats that exists on a centralized architecture or public blockchain.
- The ATR will provide full tracking of the recorded transaction for every TRUSTEE ABB that will be connected to the blockchain.
- Due to the immutability nature of blockchain, it will be impossible to alter any recorded information from the ATR.
- Hyperledger Fabric provides a unique feature called “private data collection”. The private data collection provides an extra layer of security allowing only the authorized entities to have access to the confidential transaction and also giving the option of data deletion from ATR complying with the GDPR regulation.
- The blockchain network is a decentralized distributed network; thus, there will be no central authority to drive the regular flow of the network. The network will function based on the defined policies to record transactions in the ledger.
- The main components of the blockchain network will be the ledger, **peer nodes**, ordering nodes, smart contracts, membership service provider (MSP), and so on.
- **An authorized user of other TRUSTEE components needs to be connected first to a peer node and send transaction proposal in order to write transactions to the ledger.**

ABB application view diagram

Figure 111 defines transaction recording phases at ATR. An authorized client sends a transaction proposal to a peer node for the recording of transactions. The client signature is verified, and an acknowledgement is sent without updating the ledger. A RW set of the transaction is then generated and all the endorsing peers sign the RW set. The client application submits the endorsed transactions to the orders. The ordering service nodes (OSN) combine the RW sets and endorsed transactions into blocks and send them to all the committing peers in the blockchain network. The committing peers check the new block and validate the transactions within the block. They also verify the policies, endorsements, and the accuracy of the transactions. If the transactions pass the validation checks, they are committed to the blockchain ledger in a new block. However, if the transactions do not pass the validation, they are marked as invalid transactions. After transactions are committed to the ledger, peers update the local state of the ledger.

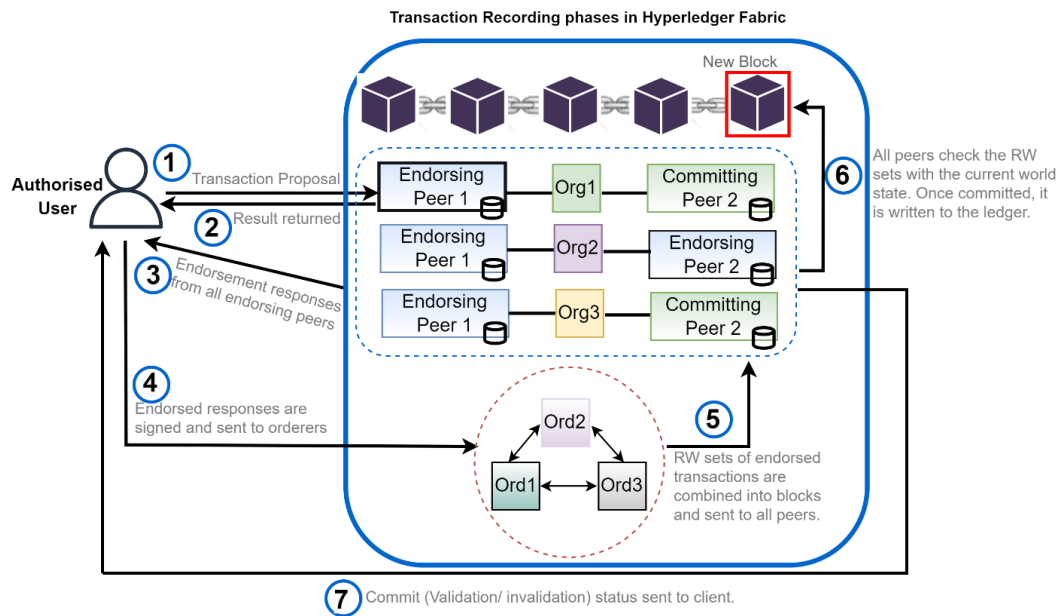


Figure 111: Transaction Recording phase at ATR

Figure 112 provides a visual representation of the ATR components and all the external interfaces associated with ATR.

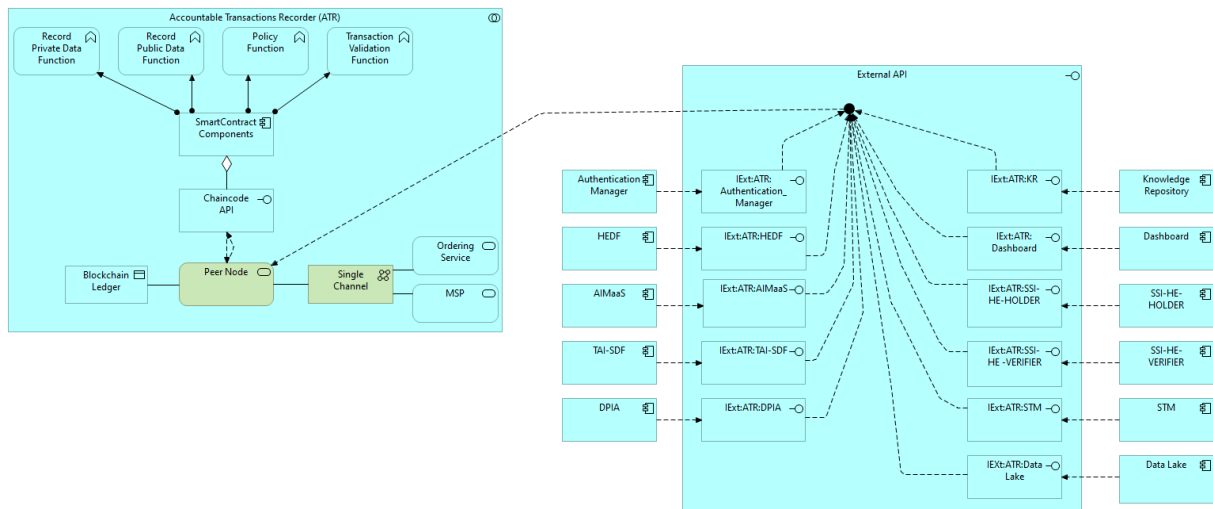


Figure 112: ATR decomposition diagram and External interfaces

ABB Sequence Diagrams

The sequence diagram presented in Figure 113 showcases the transaction recording from STM to ATR. The functionalities of ATR are then monitored.

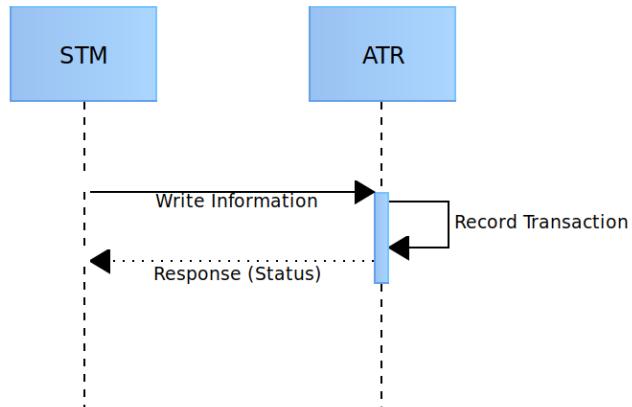


Figure 113: ATR record ABB transaction diagram (here ABB showcased is STM)

Additionally, the sequence diagram illustrating the process of a user requesting to view their history of activity (logging) from the ATR through the DA has been included in the TRUSTEE Dashboard (DA) Building Block section.

ABB External Interfaces Specification

Interface	Content	Comment
Interface ID	<i>IExt:ATR:Dashboard</i>	Any TRUSTEE component that needs to record information to ATR, will need to be an authorized user, and connected to one of the peers nodes. All the recoded transaction can be viewed by authorized entities in a Dashboard.
Interface to	DA	
Interface overview	All the recorded transaction can be accessed and viewed through the DA	

Messaging pattern	<i>Request/response, publish-subscribe, notification</i>	
Data exchanged via this interface	<i>Type and possibly structure of the data being transferred via this interface</i>	
Communication protocol	<i>e.g. gRPC, HTTP(S), REST API</i>	
Message format	<i>JSON, XML, CSV, protobuf, CBOR, BSON..</i>	
Security and integrity		

Interface	Content	Comment
Interface ID	<i>IExt:ATR:STM</i>	
Interface to	<i>Security and Trust Manager (STM) Data Protection Impact Assessment (DPIA) Knowledge Repository (KR) Data Lake (DL) HEDF AI Model as a Service (AIMaaS) SSI-HE-VERIFIER SSI-HE-HOLDER Authentication Manager (AM) TAI-SDF</i>	Many of the TRUSTEE ABBs will need to use this interface to fulfill the transaction recording process.
Interface overview	Recording transactions by receiving transaction information from the corresponding ABB	
Messaging pattern	<i>Request/response, publish-subscribe, notification</i>	
Data exchanged via this interface	<i>Transaction data</i>	
Communication protocol	<i>e.g. gRPC, HTTP(S), REST API, FABRIC API</i>	
Message format	<i>JSON, XML, CSV, protobuf, CBOR, BSON..</i>	
Security and integrity		

SSI HOMOMORPHIC FRAMEWORK BUILDING BLOCK (SSI-HE)

ABB overview

The HE Capable SSI Framework (SSI-HE) is the TRUSTEE component dedicated to the confidential characterization of datasets. This component enables the generation and issuance of verifiable credentials with different attributes that characterize a dataset. These attributes can be considered as part of the dataset identity. In TRUSTEE, attributes will be provided in a confidential way (encrypted) to avoid revealing unnecessary data but allowing confidential searches over it. For example, let's think about a data consumer needing data from Spain; each data provider will have verifiable credentials with the location attribute about his dataset. Imagine an example in which the location attribute is a specific address: "Main Street, 10, 5C, Madrid, Spain". If the attribute is shared in clear, unnecessary information is revealed such as the city, the street, the number and the floor. However, the

provision of the attribute homomorphically encrypted enhances confidentiality by allowing the search of “Spain” word inside the location attribute without revealing the rest of the location.

ABB application view diagram

For this objective, the SSI-HE is divided into three subcomponents, as shown in Figure 114:

- **Issuer:** this subcomponent provides the metadata and generates the verifiable credentials with confidential metadata. It should refer to the data owner itself (for example, in a healthcare use case, it could be the patient the data refers to) or other trusted certification entity (for example, a government, etc.).
- **Holder:** this subcomponent holds and manages the verifiable credentials with the confidential metadata. It can also verify the metadata of the dataset it deals with. It should refer to the data provider (following with the healthcare use case, it could be the hospital itself).
- **Verifier:** this component asks for some specific attribute of the metadata, received verifiable proofs of the required confidential attributes and homomorphically executes a search over the confidential attributes. It should refer to the TRUSTEE platform itself allowing a data consumer who needs datasets with specific characteristics (metadata) to filter by valid data providers. The specific characteristics of the datasets is what he searches over the encrypted attributes.

The different subcomponents interact between them for the verifiable credentials and verifiable proofs sharing, as well as for the homomorphic context and keys generation and sharing (internal interfaces will be detailed in D4.1).

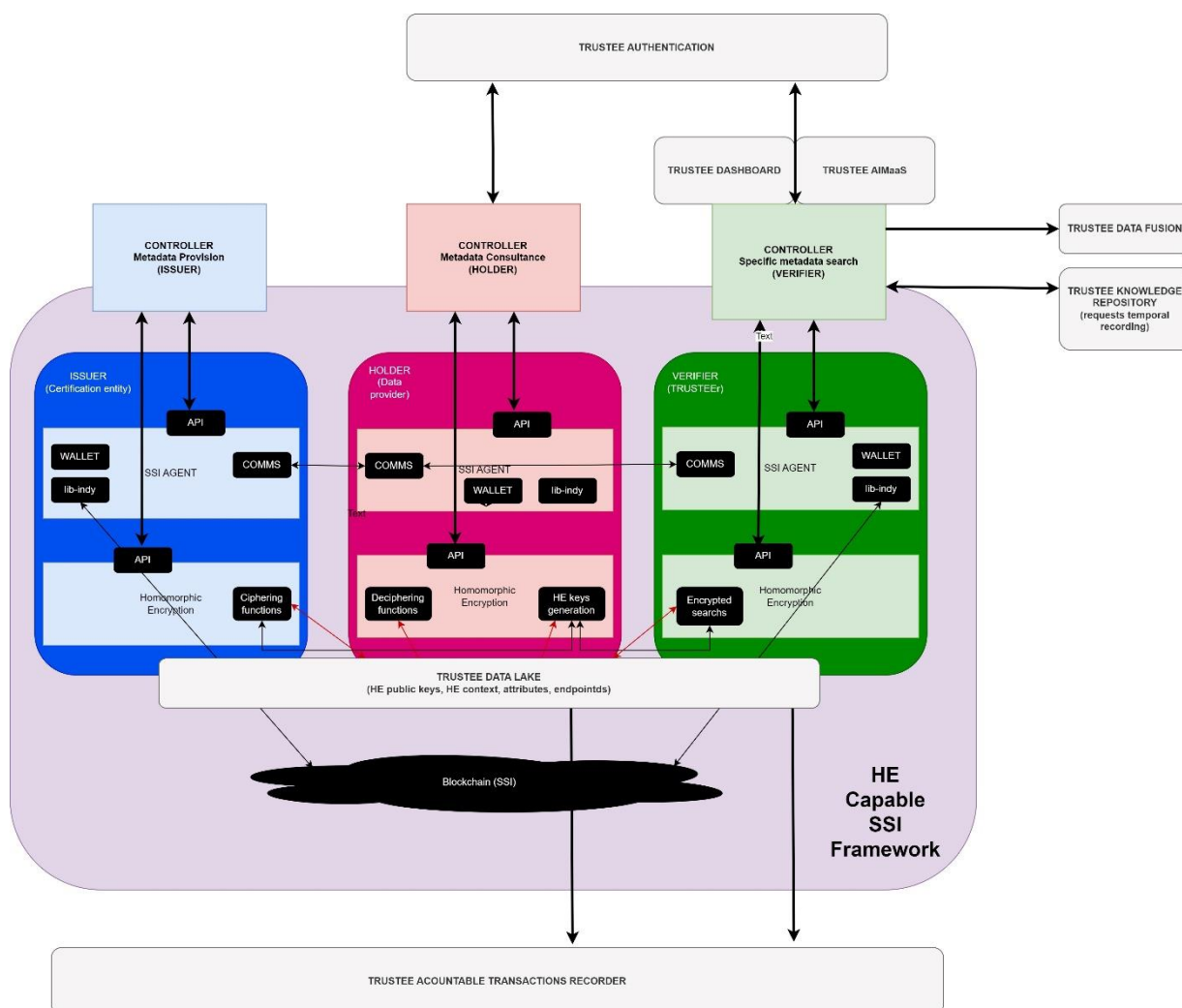


Figure 114: TRUSTEE SSI-HE Framework high level architecture

The needed external interfaces are shown in Figure 115.

- **DA:** For the graphical input of attributes to search (for example, the value Spain for location attribute) and graphical output of the obtained data providers fulfilling the requested attributes.
- **AIMaaS:** For the input of attributes to search and graphical output of the obtained data providers fulfilling the requested attributes for AI purposes (training or inference).
- **AM:** Only authenticated users can use the SSI-HE component (for data providers – holder -and data consumers – verifier -). At first, certification authorities -issuers - do not require to be part of TRUSTEE.
- **ATR:** For logging the important events of the SSI-HE operation.
- **DL:** For accessing the list of attributes available for searching functionality, the HE public keys and the endpoints of the valid data providers.
- **KR:** For temporal recording and obtaining of previous requests results (in case they are needed in a short time).
- **HEDE:** For providing the data providers endpoints fulfilling the data consumers requirements (after whole SSI-HE operation or from the KR).

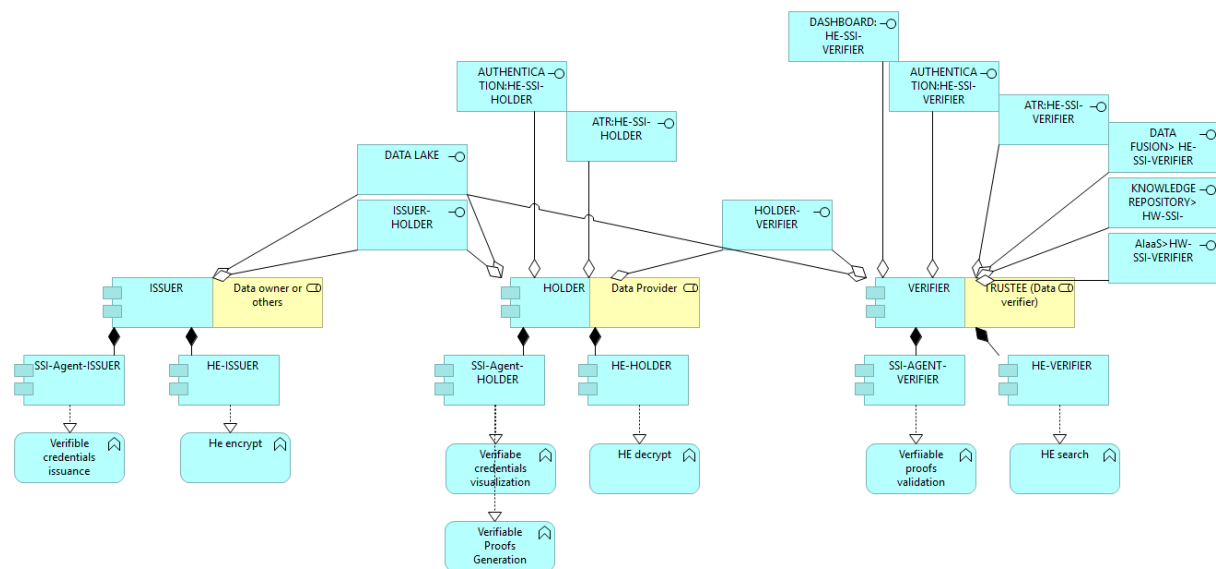


Figure 115: TRUSTEE SSI-HE Framework external interactions

ABB Sequence Diagrams

The high-level operation of this component involving other TRUSTEE components is shown in Figure 116. All interactions with data consumer will happen through the DA, having been properly authenticated in advance.

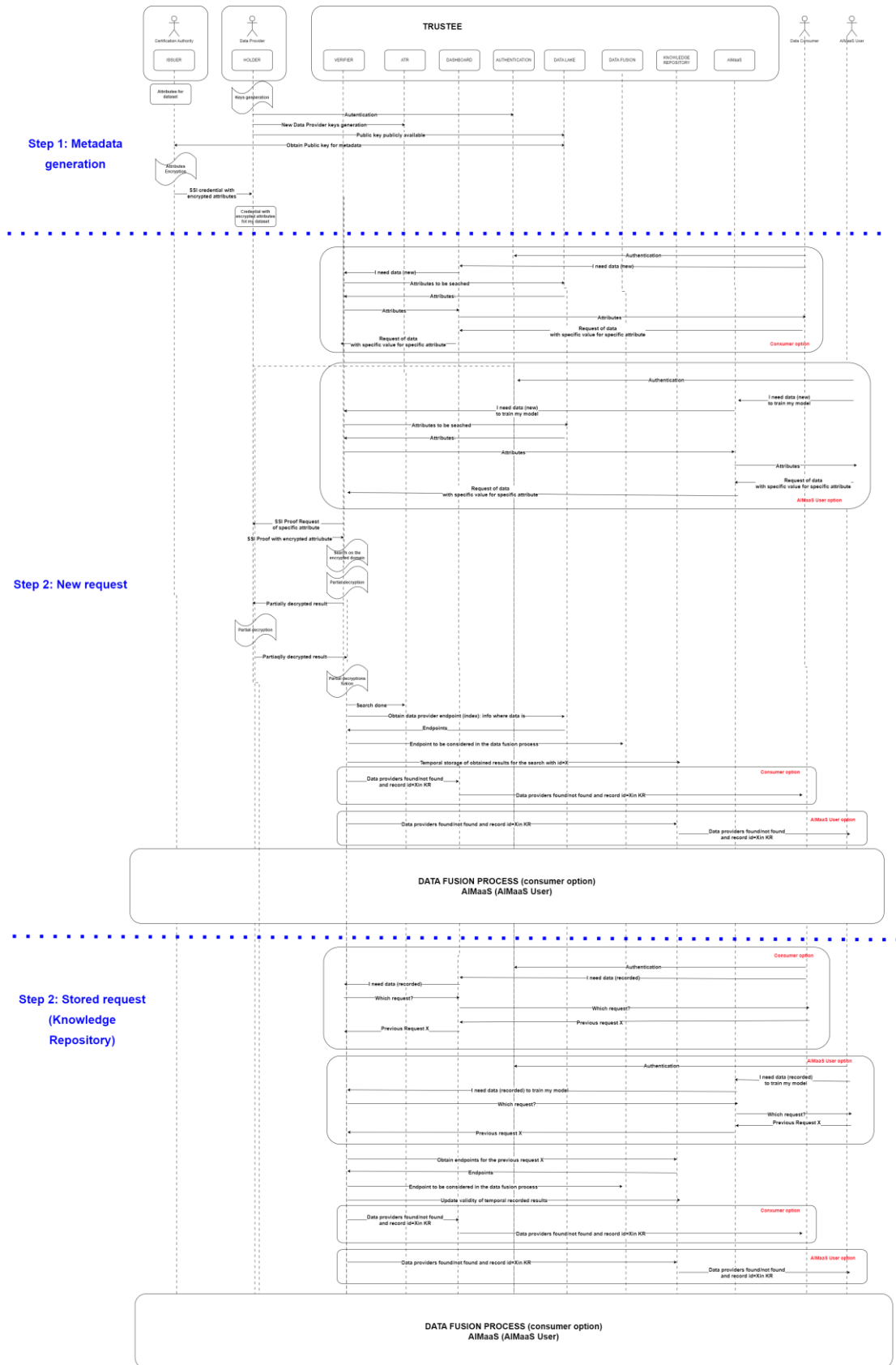


Figure 116: TRUSTEE SSI-HE Framework operation flow

ABB External interfaces specification

Issuer:

Interface	Content	Comment
Interface ID	IExt:HE-SSI-ISSUER_DL	
Interface to	Data Lake – DL	
Interface overview	Obtain the homomorphic public key for metadata	
Messaging pattern	Request/Response	
Data exchanged via this interface	Homomorphic public key for metadata	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure Communications	

Holder:

Interface	Content	Comment
Interface ID	IExt:SSI-HE-HOLDERAuthAgent:AM	
Interface to	Authentication Manager – AM	
Interface overview	Provide access to registered data providers to use the TRUSTEE platform for providing their metadata.	
Messaging pattern	Request/Response	
Data exchanged via this interface	SSI based verifiable proofs	
Communication protocol	HTTPS	
Message format	To Be Defined	Probably JSON
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:HE-SSI-HOLDER:ATR	
Interface to	Accountable Transactions Recorder – ATR	

Interface overview	Provide information about the events associated to the confidential metadata sharing (logging reasons)	
Messaging pattern	POST request	
Data exchanged via this interface	Important events	
Communication protocol	e.g. gRPC, HTTP(S), REST API	
Message format	JSON, XML, CSV, protobuf, CBOR, BSON..	
Security and integrity	Secure communications	

Interface	Content	Comment
Interface ID	IExt:HE-SSI-HOLDER:DL	
Interface to	Data Lake – DL	
Interface overview	Provide the homomorphic public key for metadata	
Messaging pattern	Request/Response	
Data exchanged via this interface	Homomorphic public key for metadata	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure Communications	

Verifier:

Interface	Content	Comment
Interface ID	IExt:SSI-HE-VERIFIER:DA	
Interface to	Dashboard – DA	
Interface overview	Provide the attribute name and value to be searched and visualize the result of the search	
Messaging pattern	Request/Response	
Data exchanged via this interface	The attribute name and value to be searched as well as the result of the search.	
Communication protocol	HTTPS	
Message format	To Be Defined	Probably JSON

Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:SSI-HE-VERIFIERAuthAgent:AM	
Interface to	Authentication Manager – AM	
Interface overview	Provide access to registered data consumers to use the TRUSTEE platform for searching for specific attributes.	
Messaging pattern	Request/Response	
Data exchanged via this interface	SSI based verifiable proofs	
Communication protocol	HTTPS	
Message format	To Be Defined	Probably JSON
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:HE-SSI-VERIFIER:ATR	
Interface to	Accountable Transactions Recorder – ATR	
Interface overview	Provide information about the events associated to the confidential metadata check: searches executed over confidential attributes (logging reasons)	
Messaging pattern	POST request	
Data exchanged via this interface	Important events	
Communication protocol	e.g. gRPC, HTTP(S), REST API	
Message format	JSON, XML, CSV, protobuf, CBOR, BSON..	
Security and integrity	Secure communications	

Interface	Content	Comment
Interface ID	IExt:HE-SSI-VERIFIER:DL	
Interface to	Data Lake – DL	

Interface overview	Obtain the available attributes/features/characteristics to be searched as well as the data providers endpoints fulfilling the attributes requirements	
Messaging pattern	Request/Response	
Data exchanged via this interface	Attributes of the datasets and Data providers endpoints	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure Communications	

Interface	Content	Comment
Interface ID	IExt:HE-SSI-VERIFIER:KR	
Interface to	Knowledge Repository – KR	
Interface overview	Store and obtain the results of previous searches (temporally stored)	
Messaging pattern	Request/Response	
Data exchanged via this interface	List of valid endpoints for a specific search	
Communication protocol	gRPC, HTTP, REST, Kafka	
Message format	JSON, protobuf, CBOR, XML, YAML	
Security and integrity	Secure Communications	

Interface	Content	Comment
Interface ID	IExt:HE-SSI-VERIFIER:HEDF	
Interface to	Homomorphic Enabled Data Fusion – HEDF	
Interface overview	Provide the suitable data provider endpoints/indexes (where datasets are accessible) to start the data fusion process	Based on the complete SSI flow or from previous queries stored on the KR.
Messaging pattern	POST Request	
Data exchanged via this interface	Endpoints/indexes	
Communication protocol	HTTPS	

Message format	JSON	
Security and integrity	Secure Communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:SSI-HE-VERIFIER:AIMaaS	
Interface to	AI Models as a Service – AIMaaS	
Interface overview	Provide the attribute name and value to be searched and visualize the result of the search	
Messaging pattern	Request/Response	
Data exchanged via this interface	The attribute name and value to be searched as well as the result of the search.	
Communication protocol	HTTPS	
Message format	To Be Defined	Probably JSON
Security and integrity	Secure Communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:AM:AuthAgent	
Interface to	Authentication Manager – AM	
Interface overview	Provide the user credentials to allow or deny access to the issuer functionality	
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	HTTP(S) / OIDC	
Message format	JSON	
Security and integrity	Secure Communications (HTTPS)	

Exposed external API definition of SSI-HE components

There has been a preliminary API definition of SSI-HE components which will be further investigated and discussed in upcoming versions of this deliverable, namely D2.2 and D2.3. Endpoints required for interactions between issuer, holder, and verifier have not been defined yet; only endpoints required for external interactions – physical actors or other TRUSTEE components, mainly, DA.

SSI Issuer

Issue_credentials (data provider id, attribute[], values [])

- *Parameters:* data provider id, list of attributes, value of these attributes
- *Output:* the output will be a true (operation correctly done) or false (error).
- *Expected executor:* Certification Authority through a graphical interface.
- *Context:* the certification authority wants to issue a credential to a data provider Id certifying specific attributes/values of the data provider's datasets.

SSI Holder

HE_keys_generation()

- *Parameters:* No parameter is needed
- *Output:* the output will be a true (operation correctly done) or false (error).
- *Expected executor:* Data provider through a graphical interface
- *Context:* the data provider requests the HE keys generation.

SSI Verifier

Data_Request (type)

- *Parameters:* type of request: new (complete SSI process) or recorded (from KR)
- *Output:* the output will be a list of attributes to be searched (if new request) or a list of recorded previously requests (if recorded type).
- *Expected executor:* Data consumer through the DA or AI developer through AIMaaS.
- *Context:* the data consumer makes a request of data to be processed.

New_Request_Definition (attribute[], value[])

- *Parameters:* list of attributes and corresponding values to use for searching valid data providers.
- *Output:* the output will be a true (data providers found) or false (data providers not found) as well as the id of the request stored in the KR for future requests.
- *Expected executor:* Data consumer through the DA or AI developer through AIMaaS.
- *Context:* the data consumer defines the required attributes of the data he needs.

Stored_Request_Definition (id)

- *Parameters:* id of the desired request recorded on the KR
- *Output:* the output will be a true (data providers found) or false (data providers not found).
- *Expected executor:* Data consumer through the DA or AI developer through AIMaaS.
- *Context:* the data consumer defines the previous request to be performed.

TRUSTWORTHINESS AI SUPPORT DESIGN FRAMEWORK (TAI-SDF) BUILDING BLOCK

ABB overview

The Framework takes into account a typical assessment and system monitoring audit workflow to develop explainable and trustworthy AI for efficient and robust use and re-use of data across interdisciplinary domains (Energy, Health, Education, Automotive, Space):

- Assess the conformance of the target object in scope using the Compliance Support Tool
- Adherence to ethical principles based on fundamental rights (e.g., fairness and explainability)
- Risk analysis considering safety, security, and privacy
- Continuous monitoring across each phase of the AI system lifecycle
- Evaluation of AI robustness
- Adoption of recent programming practices and innovative secure deployment methods

ABB application view diagram

The TAI-SDF architecture consists of several modules Administration module, Authentication module, HomePageModule, OperationModule, SurveyModule, and the GDPR Compliance checker. The overall architecture with its external interfaces is depicted in Figure 118 below.

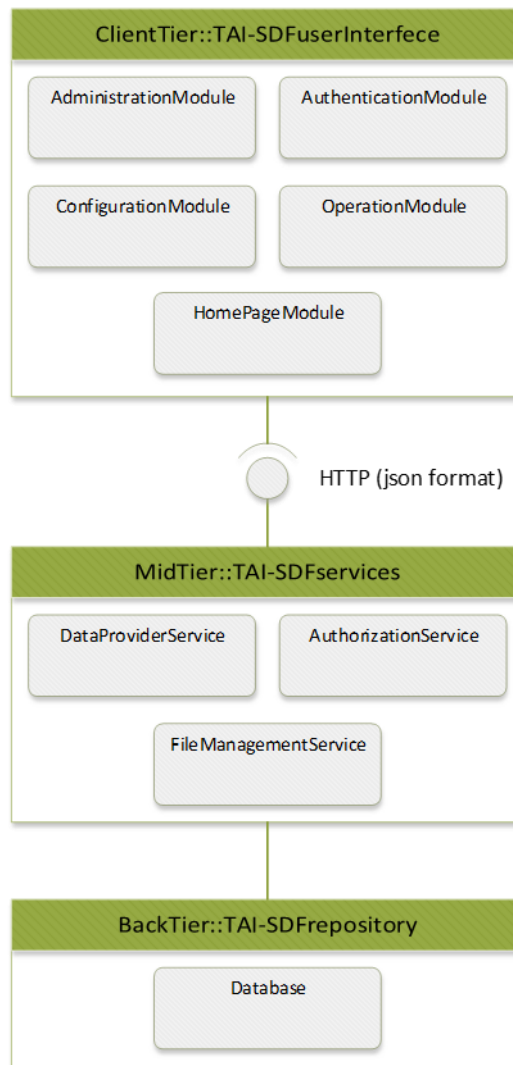


Figure 117: TAI-SDF

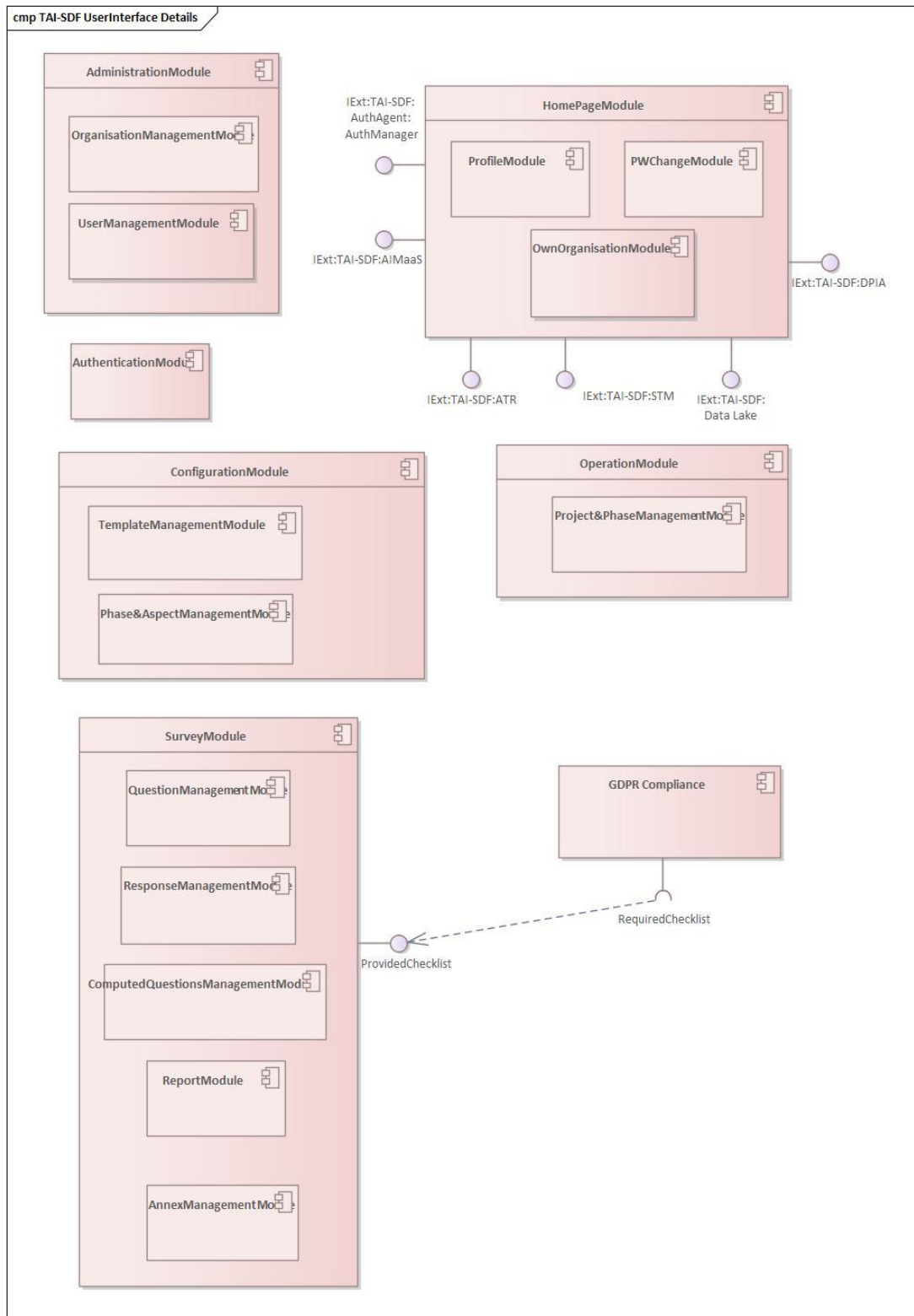
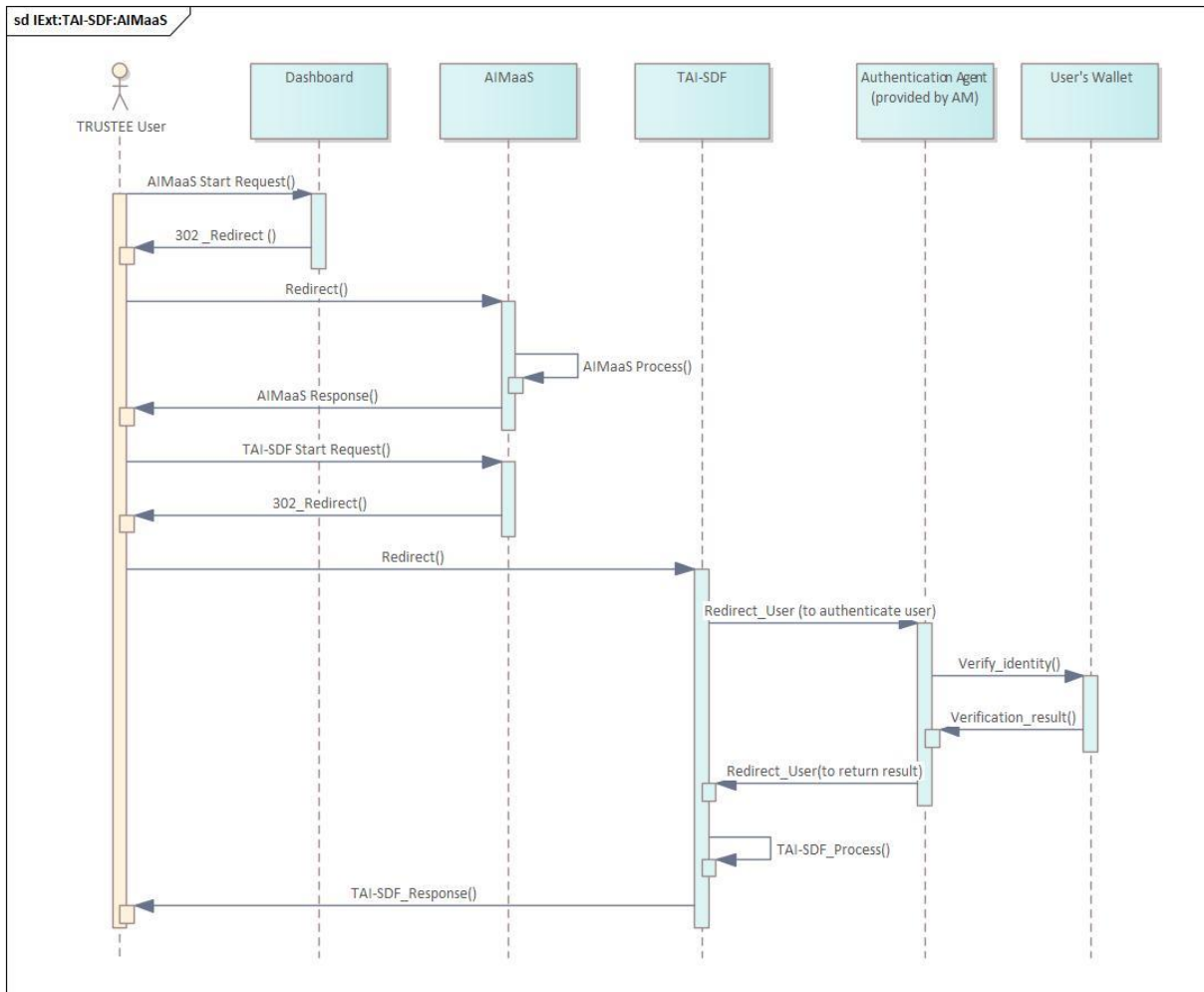


Figure 118: TAI-SDF architecture with inner modules and external interfaces

As mentioned earlier, the TAI-SDF architecture is composed by the following modules:

- **ConfigurationModule** that contains all the components, functions and procedures to define matrixes and templates;
- **HomePageModule** that contains all the components, functions a procedures to access to organization and profile management capabilities and to the external interface of the TAI-SDF;
- **OperationalModule** that contains all the components, functions and procedures to create, modify and execute projects;
- **AdministrationModule** that contains all the components, functions and procedures to manage organization and users;
- **SurveyModule** that contains all the components, function and procedure to define and execute questionnaires and tests;
- **GDPR Compliance** that contains all the functions and procedures to manage the GDPR checklist;
- **AuthenticationModule** that contains all the functions and procedures to interface with the authentication agent.

ABB Sequence diagrams



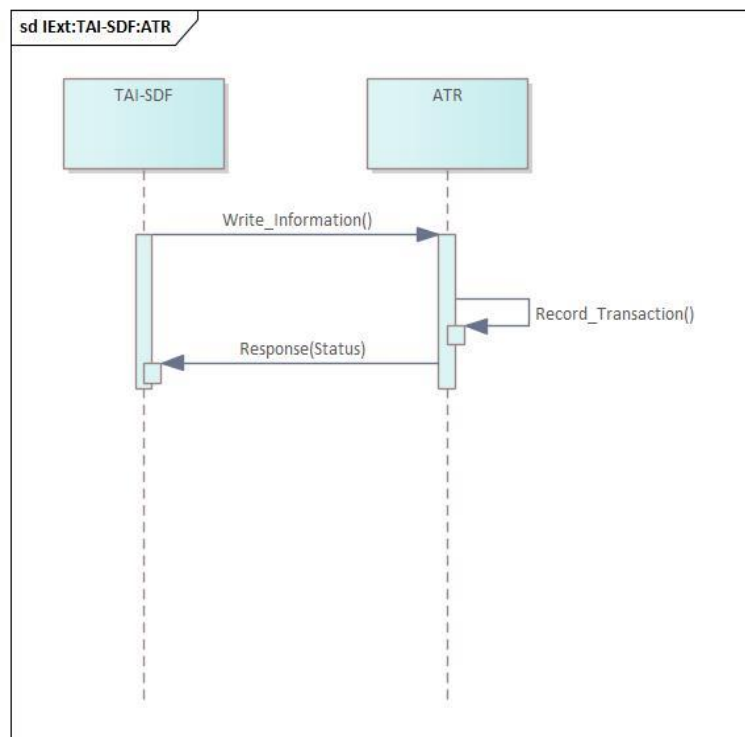
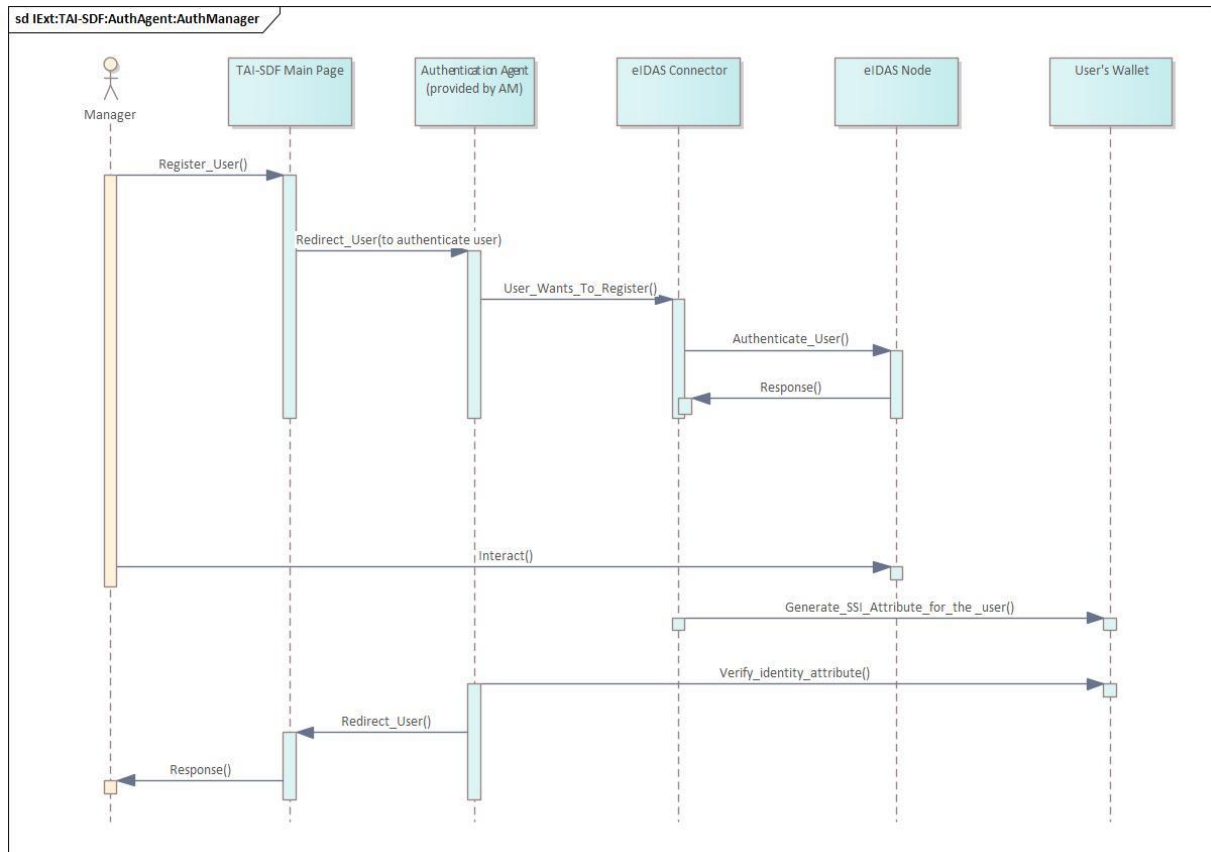


ABB External interfaces specification

Interface	Content	Comment
Interface ID	IExt:TAI-SDF:STM	
Interface to	Security and Trust Manager – STM	
Interface overview	Purpose of the interface is to exchange checklist results in json format produced by GDPR Compliance checker	
Messaging pattern	Request/response	
Data exchanged via this interface	Questions and answers related to the GDPR assessment	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	The data transmission should be secured.	

Interface	Content	Comment
Interface ID	IExt:TAI-SDF:AIMaaS	
Interface to	AI Models as a Service – AIMaaS	
Interface overview	The user should use the AIMaaS and call TAI-SDF	
Messaging pattern	Request/Response	
Data exchanged via this interface	Assessment report will be sent back to AIMaaS	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	The data transmission should be secured	

Interface	Content	Comment
Interface ID	IExt:TAI-SDFAuthAgent:AM	
Interface to	Authentication Manager – AM	Many of the TRUSTEE ABBs will need to use this interface to authenticate users.
Interface overview	The interface is associated to an Authentication Agent, a software will be part of the TAI-SDF which will require to verify that a user is already authenticated or requires to be authenticated before accessing the service, or to authenticate a user.	
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	<i>HTTP(S) / OIDC</i>	
Message format	<i>JSON</i>	
Security and integrity	The data transmission should be secured.	

Interface	Content	Comment
Interface ID	IExt:TAI-SDF:ATR	
Interface to	Accountable Transactions Recorder – ATR	
Interface overview	Activities of TAI-SDF that should be recorded to ATR: • Add Answers to questions • Modify answers to questions • Delete answers to questions	
Messaging pattern	Publish/subscribe	

Data exchanged via this interface	<i>A record: <userID, timestamps, actionType></i>	
Communication protocol	FABRIC API	
Message format	protobuf	
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	

Interface	Content	Comment
Interface ID	IExt:TAI-SDF:DL	
Interface to	Data Lake – DL	
Interface overview	The goal is to provide the store the users' answers to the questionnaires	
Messaging pattern	To be defined	
Data exchanged via this interface	Publish/subscribe	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	The data transmission should be secured.	

Interface	Content	Comment
Interface ID	Iext:TAI-SDF:DPIA	
Interface to	Data Privacy Impact Assessment – DPIA	
Interface overview	The goal is to get the results from TRUSTEE PIA tool in the privacy module	
Messaging pattern	PIA report	

Data exchanged via this interface	Request/Response	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	The data transmission should be secured.	

AI MODEL AS A SERVICE (AIMAAS) BUILDING BLOCK

ABB overview

The AI Models as a Service building block will enable a plug-n-play approach for the on-boarding of various AI models for data processing. This block will offer functionalities for both Consumers, as well as Model Providers. Consumers will be able to launch the AIMaaS tool directly through the DA ABB, search for datasets of their choice using the SSI-HE framework, search for available AI models, and run a specific AI model on specific data, as a service. On the other hand, the Model Providers will be able to upload to the TRUSTEE platform their pre-trained models, they will have the opportunity to run explainable functions to their trained models and perform a trustworthiness assessment of the model through the TAI-SDF ABB. Finally, they will be able to store their models within the TRUSTEE environment in a form that will enable Consumers to use them in an easy and fast mode.

Besides simple AI/ML models that will be included, the AIMaaS will also offer FL models for small fast analytics need in the source.

The ABBs that interact with the AIMaaS block, through the defined external interfaces , are the following:

- **DA** (IExt:AIMaaS:DA). The AIMaaS tool will be directly initiated by the user (either Consumer or Model Provider) through the initial point of entrance to the TRUSTEE platform, i.e., the Dashboard.
- **AIMaaS**: For the input of attributes to search and graphical output of the obtained data providers fulfilling the requested attributes for AI purposes (training or inference).
- **AM** (IExt:AIMaaS:AuthAgent:AM). The AIMaaS user will need to verify their authentication with the AM ABB, as only authenticated users should be able to use the TRUSTEE tools.
- **SSI-HE Verifier** (IExt:AIMaaS:SSI-HE-Verifier). The user will be able to search for datasets by attribute, using the functionalities provided by the SSI-HE framework.
- **HEDF** (IExt:AIMaaS:HEDF). The homomorphically encrypted data that each user chooses to use will be transferred to the HEDF block and all computations on them will take place within this ABB, including training of an AI model, as well as running of a model.
- **DL** (IExt:AIMaaS:DL). The AI models will be stored in the DL, so the data consumers will be able to search and choose a model of their choice, while the model providers will be able to store there all the necessary files to deploy and run their model as a service.
- **ATR** (IExt:AIMaaS:ATR). All the transactions performed by any user will be recorded in the ATR.
- **KR** (IExt:AIMaaS:KR). The result of all the inference runs of the AI models will be stored in the knowledge repository (if the Model Provider and potential Data Provider have agreed on this), so that if a future user would want to run the same model on the same data, the result would be retrieved from KR, thus speeding up the process.

- **TAI-SDF** (IExt:AIMaaS:TAI-SDF). The model providers who want to offer their models to the TRUSTEE platform will be able to perform a trustworthiness assessment of their models and receive a respective report.
- **OneSS** (IExt:AIMaaS:OneSS). The model provider will need to access the TRUSTEE wiki, which is part of the OneSS. Thus, they will be guided to follow all the necessary steps, in order to upload their model in the appropriate form, for the consumers to be able to run them. Additionally, several functionalities of the AIMaaS ABB will be exposed via APIs through the OneSS for the Developer Persona to use as part of the incremental strategy of TRUSTEE; However, in this case no data of the Data Providers will be made available and/or AI/ML models trained with the data of the Data Providers. The functionalities to be offered through OneSS will be further investigated in the upcoming months of the project and will be further discussed in next versions of this deliverable, namely D2.2 and D2.3 as well as in upcoming deliverables of WP3 and WP4.

The explainable AI functions will be incorporated into the AIMaaS block, so no external interface was defined. The ABB application view diagram, the sequence diagrams that correspond to the flow of a Consumer running a model and the flow of a Model Provider training and storing their model, as well as the external interfaces definition, follow.

ABB application view diagram

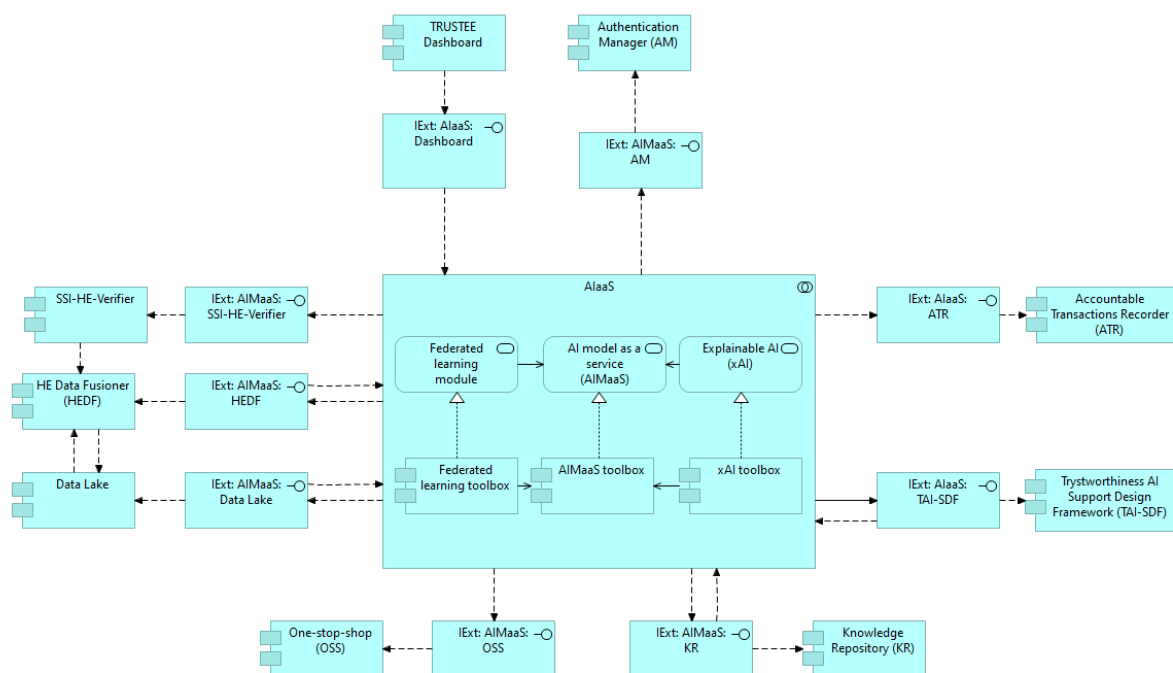


Figure 119: AIMaaS application view diagram

ABB Sequence Diagrams

The sequence diagram of the Consumer flow (Figure 120) shows an initial version of the steps that will be performed within the TRUSTEE platform, in the case a Consumer wants to run a pre-trained model, that is stored within TRUSTEE, on a dataset provided by a Data Provider. Following, the Consumer will enter the TRUSTEE environment through the DA, where they will be redirected to the AIMaaS module. Then, the AIMaaS will check the authentication of the Consumer, through the AM.

Afterwards, the Consumer searches for data sets by attribute through the SSI-HE. The steps shown in Figure 120 will be followed by the SSI-HE building block (which have been omitted in the current sequence diagram for convenience), and in the case a suitable dataset has been found, a corresponding message will be sent back to the Consumer. Additionally, the dataset index will be retrieved from the DL and sent to the AIMaaS, as well as to the HEDF, for the data fusion process to take place.

The next step involves searching for a model in the DL. The data consumer will search for a suitable pre-trained model, based on the model description and the HE schema that has been adopted during the training procedure. If a suitable model is found, a corresponding message, along with the model index will be sent to the AIMaaS.

Subsequently, the AIMaaS will send the data and model index to the DL, which will search in the corresponding log for previous runs of this model on the same dataset. The entries of this log will contain a model index and a dataset index, showing that an inference request for this model on this dataset has been performed in the past, and an index of the KR, pinpointing to the location where the result of this previous inference request is stored. If such previous request exists, then the KR index is returned to the AIMaaS, which requests the result from the KR, which in turn sends the result back to the AIMaaS. This way we take advantage of previous computations, in order to reduce the computational cost of running an extra inference on the model.

If no previous runs of this model have taken place, the AIMaaS sends the model index to the HEDF, which has to request the model from the DL, where it is stored. The model is sent to the HEDF, which runs it on the chosen data and sends the inference result back to the AIMaaS. Then, the AIMaaS sends the model index, the dataset index, as well as the inference result, to the KR for future use, if the Model Provider and potential Data Provider have agreed on this. The KR will need to send the corresponding KR index where this information will be stored, to the DL, so that it can update the log with a new entry, containing the information of the current inference.

After every action that is performed by the AIMaaS, the corresponding information is sent to the ATR, so that it can record the related transaction.

The steps described above are subject to changes and updates in upcoming versions of this deliverable, namely D2.2 and D2.3.

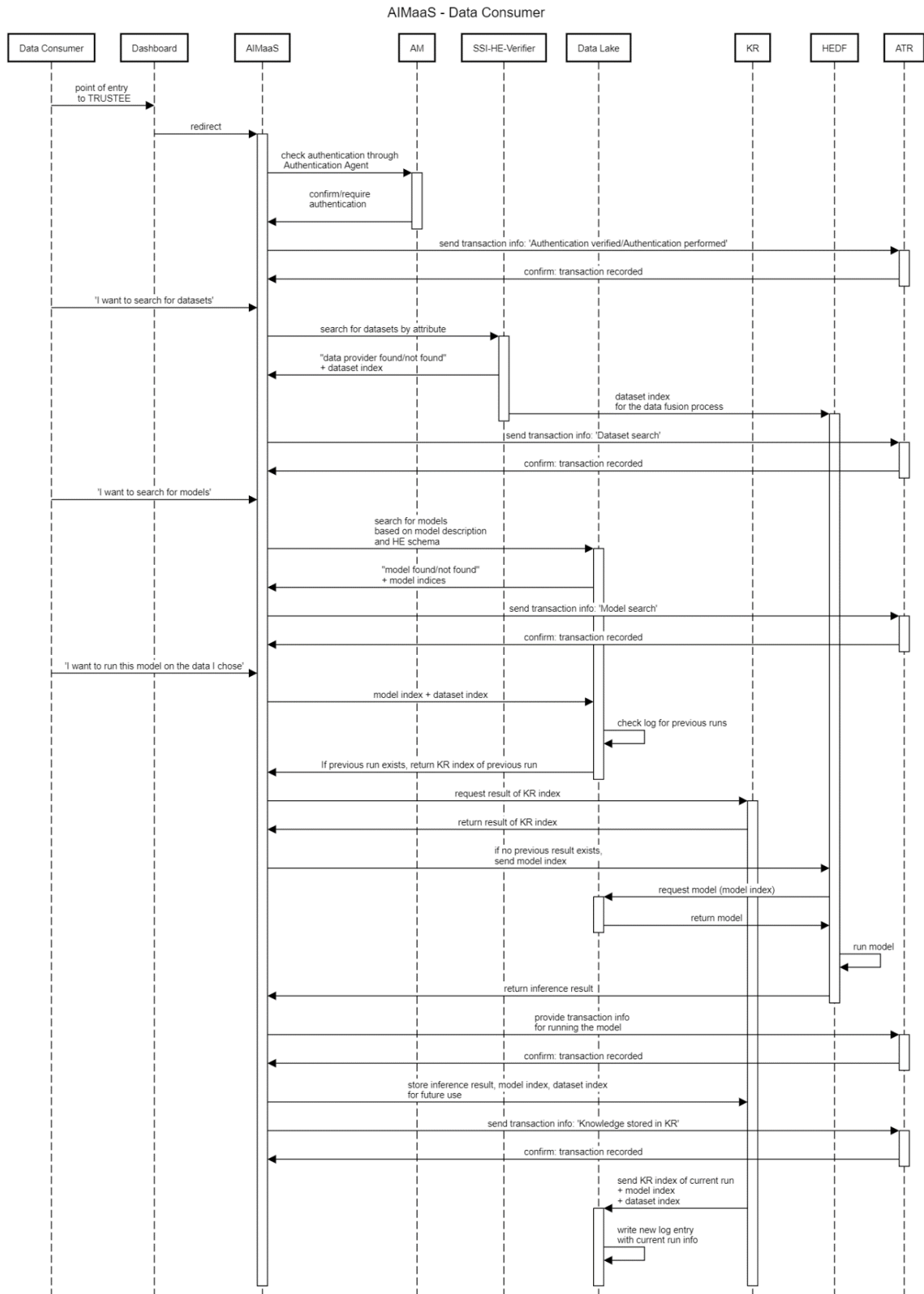


Figure 120: Sequence diagram of the use case of a Data Consumer using the AIMaaS

The Model Provider sequence diagram (Figure 121) showcases the flow of a Model Provider who wishes to upload a self-developed and pre-trained model on the TRUSTEE Platform. As in the Consumer flow, the Model Provider will enter the TRUSTEE environment through the DA, be redirected to the AIMaaS ABB and their authentication will be checked by the AM. Afterwards, the model provider needs to access the TRUSTEE wiki that resides in the OneSS, in order to check the steps that need to be followed. The first step includes running an xAI function on the trained model. Then, the Model Provider can start the TAI-SDF module, so they can run a trustworthiness assessment on the model and get a corresponding report back.

It is the Model Provider's responsibility to create all the necessary files in order to run the model in a container. Then, the Model Provider needs to send the necessary files (including the docker image, the model description, the trustworthiness assessment report, etc.) to the DL, where they will be stored. The model index of this newly stored model can be sent back to the Model Provider for future use.

Again, as in the case of the data consumer flow, after every action that is performed, the corresponding information is sent to the ATR, so that it can record the related transaction.

The steps described above are subject to changes and updates in upcoming versions of this deliverable, namely D2.2 and D2.3.

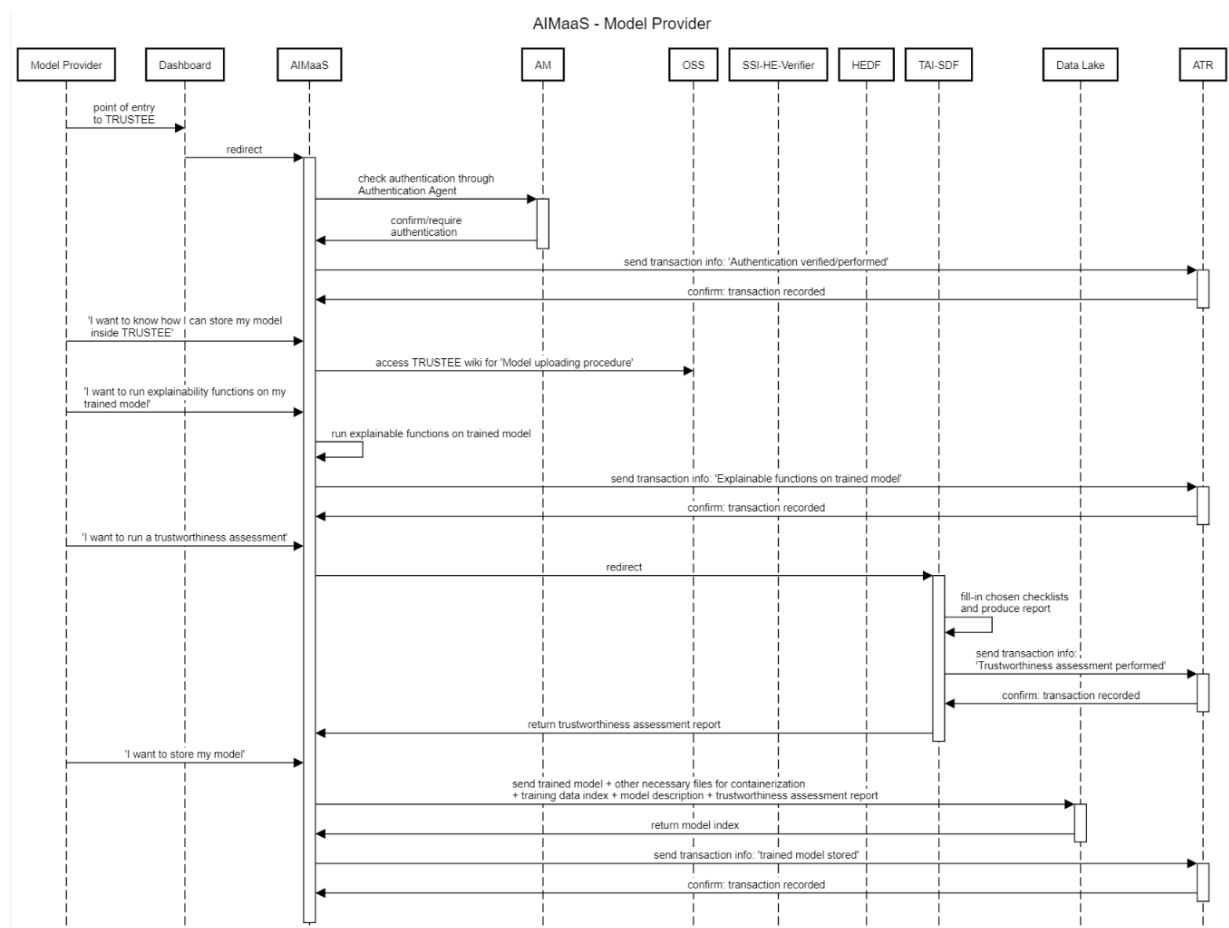


Figure 121: Sequence diagram of the use case of a Model Provider using the AIMaaS

ABB External interfaces specification

Interface	Content	Comment
Interface ID	IExt:AIMaaS:DA	
Interface to	Dashboard – DA	
Interface overview	A TRUSTEE user will be able to launch the AIMaaS tool through the DA	
Messaging pattern	Redirect	
Data exchanged via this interface	A ping to start the AIMaaS tool	
Communication protocol	HTTP(S)	
Message format	URL query	
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:AIMaaS:TAI-SDF	
Interface to	Trustworthy AI Support Design Framework – TAI-SDF	
Interface overview	The TRUSTEE user will be able to launch the TAI-SDF tool directly from the AIMaaS	
Messaging pattern	redirect	
Data exchanged via this interface	A ping to launch the TAI-SDF, trustworthiness assessment report	
Communication protocol	HTTPS	
Message format	URL query, any text-like format	
Security and integrity	Secure communications (HTTPS)	The data transmission should be secured.

Interface	Content	Comment
Interface ID	IExt:AIMaaS:ATR	
Interface to	Accountable Transactions Recorder – ATR	

Interface overview	Activities of AIMaaS that should be recorded to ATR: • Authentication of user • Query for datasets • Query for data models • Run a model (consumer) • Upload a model, ready to be deployed (model provider) • Train a model (model provider) • Store extracted knowledge	
Messaging pattern	Publish/subscribe	
Data exchanged via this interface	Message exchanged: <user ID, type of action, timestamp, other info> Other info includes other files needed for specific transactions (i.e. model index, dataset index, etc.)	
Communication protocol	FABRIC API	
Message format	protobuf	
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	The data transmission should be secured.

Interface	Content	Comment
Interface ID	IExt:AIMaaS:HEDF	
Interface to	Homomorphic Enabled Data Fusion – HEDF	
Interface overview	The AIMaaS models will run within the HE Data Fusioner	
Messaging pattern	Request/response	
Data exchanged via this interface	ID of the data that should be accessed by the AIMaaS	
Communication protocol	HTTPS	
Message format	json	
Security and integrity	Secure communications (HTTPS)	The data transmission should be secured.

Interface	Content	Comment
Interface ID	IExt:AIMaaS: DL	
Interface to	Data Lake – DL	
Interface overview	The AIMaaS will search for models, will request models and will store models in the DL	
Messaging pattern	Request/response	
Data exchanged via this interface	AI models, files necessary for the deployment of models, model description, model index, dataset index	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure communications (HTTPS)	The data transmission should be secured.

Interface	Content	Comment
Interface ID	IExt:AIMaaS:SSI-HE-VERIFIER	
Interface to	SSI Homomorphic Capable Framework – SSI-HE-VERIFIER	
Interface overview	The TRUSTEE user will provide to the SSI-HE-VERIFIER the attribute name and value to be searched and visualize the result of the search through the AIMaaS	
Messaging pattern	Request/response	
Data exchanged via this interface	The attribute name and value to be searched, the result of the search.	
Communication protocol	HTTPS	
Message format	json	
Security and integrity	Secure communications (HTTPS)	The data transmission should be secured.

Interface	Content	Comment
Interface ID	IExt:AIMaaS:KR	

Interface to	Knowledge repository – KR	
Interface overview	Knowledge resulting from the AIMaaS will be stored in the Knowledge Repository	
Messaging pattern	Request/response	
Data exchanged via this interface	Model inference result, model index, dataset index, timestamp	
Communication protocol	IPFS API	
Message format	json	
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	The data transmission should be secured.

Interface	Content	Comment
Interface ID	IExt:AIMaaSAuthAgent:AM	
Interface to	Authentication Manager – AM	<i>Many of the TRUSTEE ABBs will need to use this interface to authenticate users.</i>
Interface overview	The interface is associated to an Authentication Agent that as a software will be part of the AIMaaS which will require to verify that a user is already authenticated or requires to be authenticated before accessing the service, or to authenticate a user.	
Messaging pattern	Request/response	
Data exchanged via this interface	The needed attributes from the user's identity.	
Communication protocol	HTTP(S) / OIDC	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:AIMaaS:OneSS	
Interface to	One-Stop-Shop – OneSS	

Interface overview	AIMaaS will need access to the OneSS wiki, so that model providers are guided through the process of model training and uploading.	
Messaging pattern	Request/response	
Data exchanged via this interface	-	
Communication protocol	HTTPS	
Message format	json (if data exchange is needed)	
Security and integrity	Secure communications (HTTPS)	The data transmission should be secured.

DATA LAKE (DL) BUILDING BLOCK

ABB overview

The TRUSTEE DL will be a solution that will centrally store the:

- public HE keys,
- dataset indexes,
- datasets and attributes catalogs,
- AI/ML models & algorithms,
- logs with information on models' executions,
- semantic ontologies that will describe the datasets' schemas.

The TRUSTEE DL has a central storage space where it stores all the above., while the data sources are distributed and data always reside at the providers' side while the DL includes the endpoints for these datasets.

Taking into consideration that Consumers that will use the TRUSTEE solution will provide characteristics of the data they require for their purposes, the DL will provide the storage indexes of the Data Providers' data that fit their requirements. On the Data Providers' side, each Data Provider will also have decentralized data lakes on their own premises for storing their own data.

The DL interacts with the following TRUSTEE ABBs and external interfaces:

1. DEDF ABB (IExt:DL:HEDF). The DL will store the indexes for the datasets that are available from each data provider.
2. SSI-HE Holder sub-component (IExt:DL:SSI-HE-HOLDER). The DL stores and provides when requested the data providers' public keys that are generated. Their communication ensures secure metadata exchange and management.
3. SSI-HE Issuer sub-component (IExt:DL:SSI-HE-ISSUER). The DL stores and provides when requested the homomorphic public keys for metadata. Their communication ensures secure metadata exchange.
4. SSI-HE Verifier sub-component (IExt:DL:SSI-HE-VERIFIER). The DL provides the data provider endpoints / indexes, i.e., information where the data is, to the VERIFIER.
5. ATR (IExt:DL:ATR). The DL will send to the ATR any transaction records carried out.

6. External connectors of the OneSS ABB (IExt:DL:OneSSExternal Connectors). The DL will store External data sources' data information for indexing. This will be further investigated in upcoming months of the project and will be further discussed in upcoming versions of this deliverable, namely D2.2 and D2.3 and in upcoming deliverables of WP4.
7. AM (IExt:DLAuthAgent:AM). DL and AM exchange any required user authentication information.
8. AIMaaS (IExt:DL:AIMaaS). The DL will provide to the AIMaaS ABB the KR index for any existing model runs for which the KR has information on and will receive new AI models and logs for storage. AIMaaS will send to DL all trained AI model information (trained model results, datasets indexes used for the training etc.), once the AIMaaS retrieves additional info from the Data Provider (model description, files necessary for the deployment of the model, the trustworthiness assessment report etc.)
9. KR (IExt:DL:KR). The DL will search the KR for any previous logs for previous AI models runs (based on the AI model and dataset indexes). If previous run exists, the DL gets the KR index to share to the AIMaaS. KR will send to the DL the index of the last entry run performed by AIMaaS, along with the model and dataset indexes, which will record a new entry in the log with the model, dataset, and KR index of the current run

ABB application view diagram

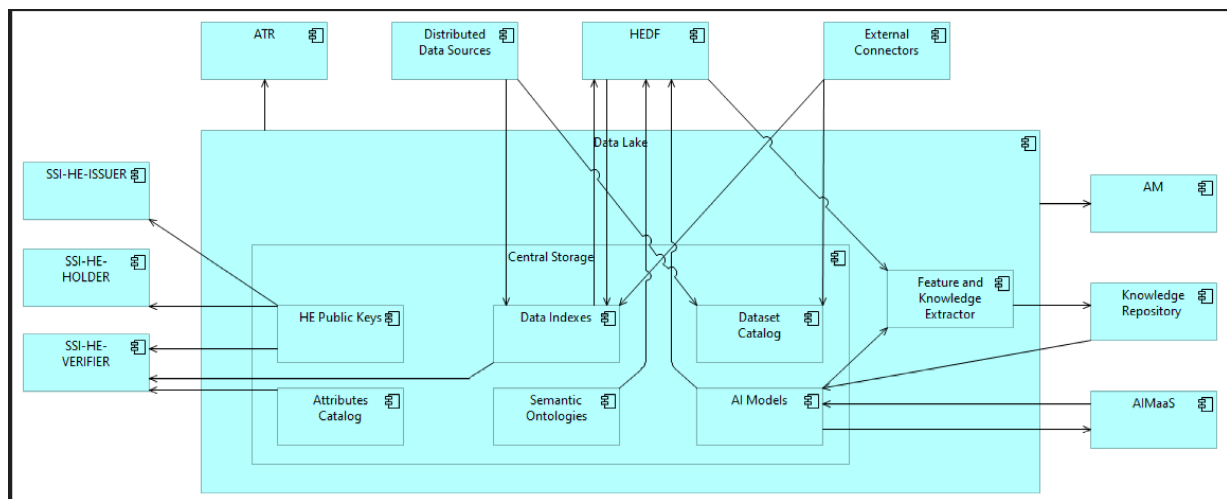


Figure 122: DL ABB breakdown diagram

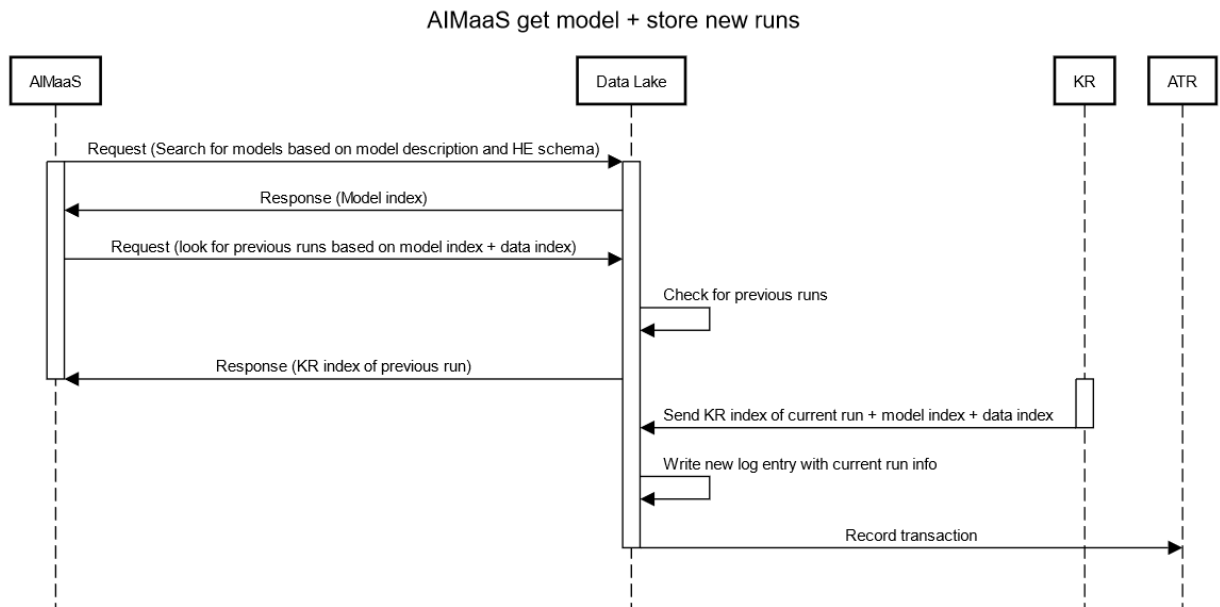


Figure 123: Sequence diagram: DL – AIMaaS (get model & store new runs)

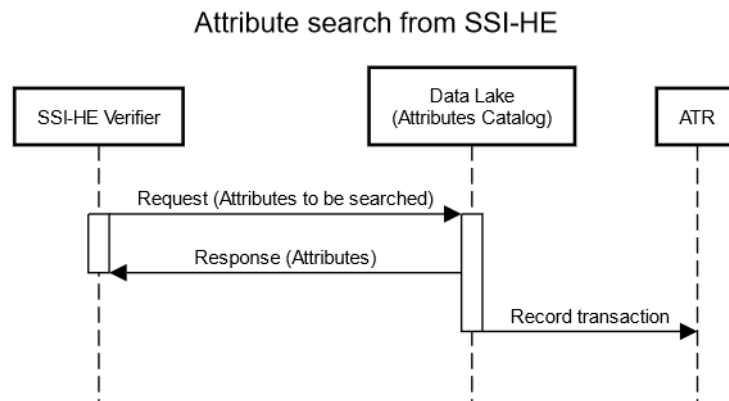


Figure 124: Sequence diagram: DL – SSI-HE (search attributes)

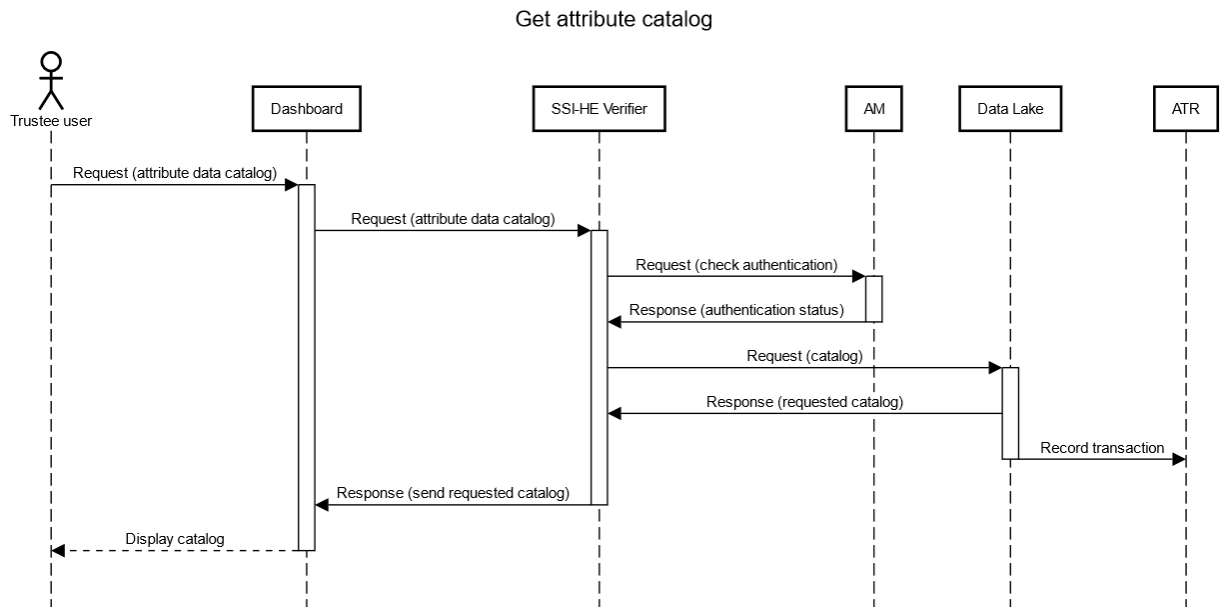


Figure 125: Sequence diagram: DL – SSI-HE (get attribute catalog)

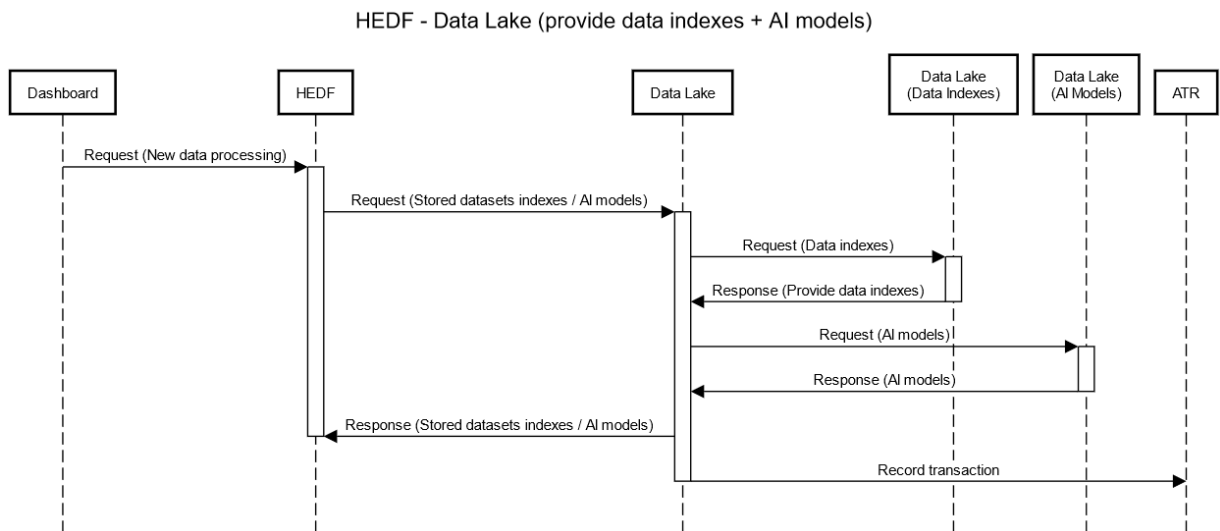


Figure 126: Sequence diagram: DL – HEDF (provide indexes & models)

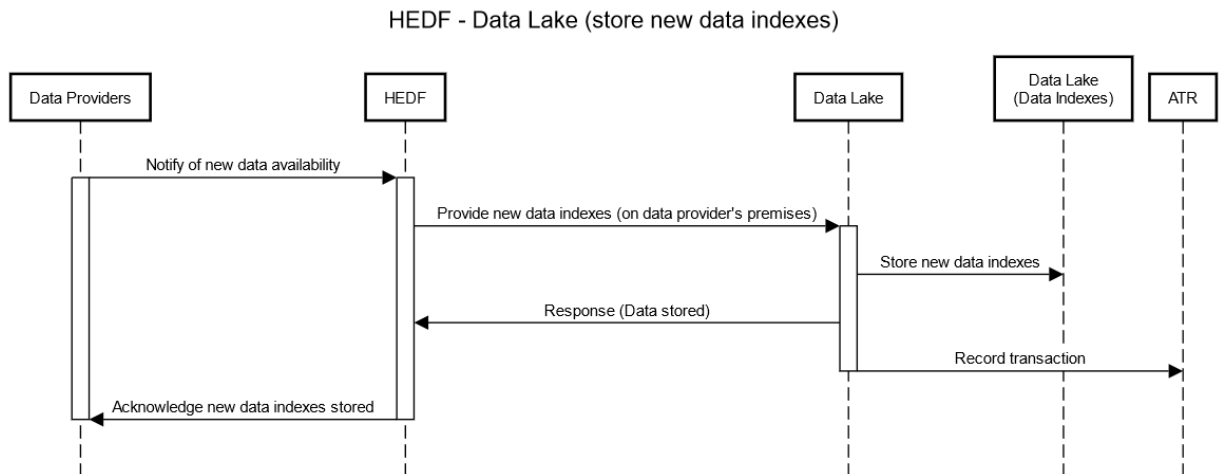


Figure 127: Sequence diagram: DL – HEDF (store new indexes)

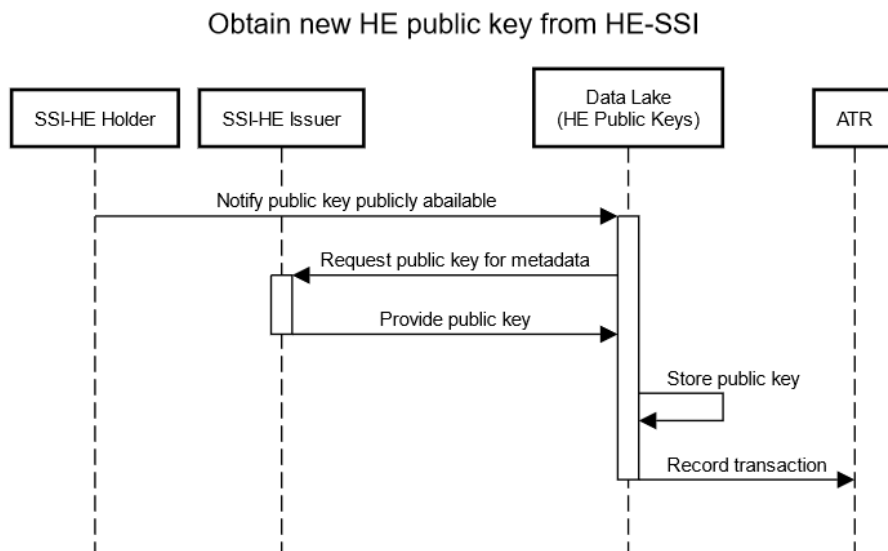


Figure 128: Sequence diagram: DL – SSI-HE, obtaining public keys

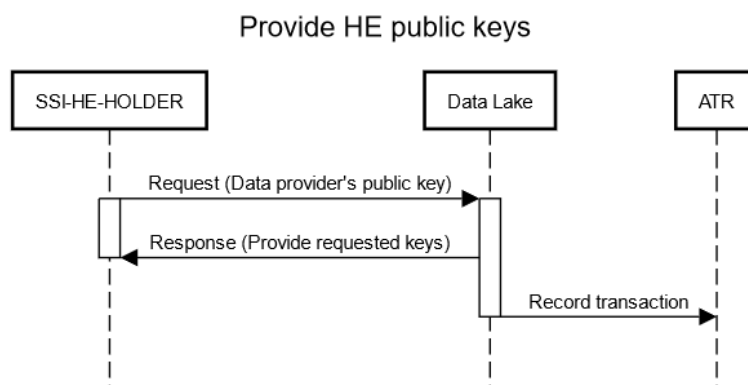


Figure 129: Sequence diagram: DL – SSI-HE, providing public keys

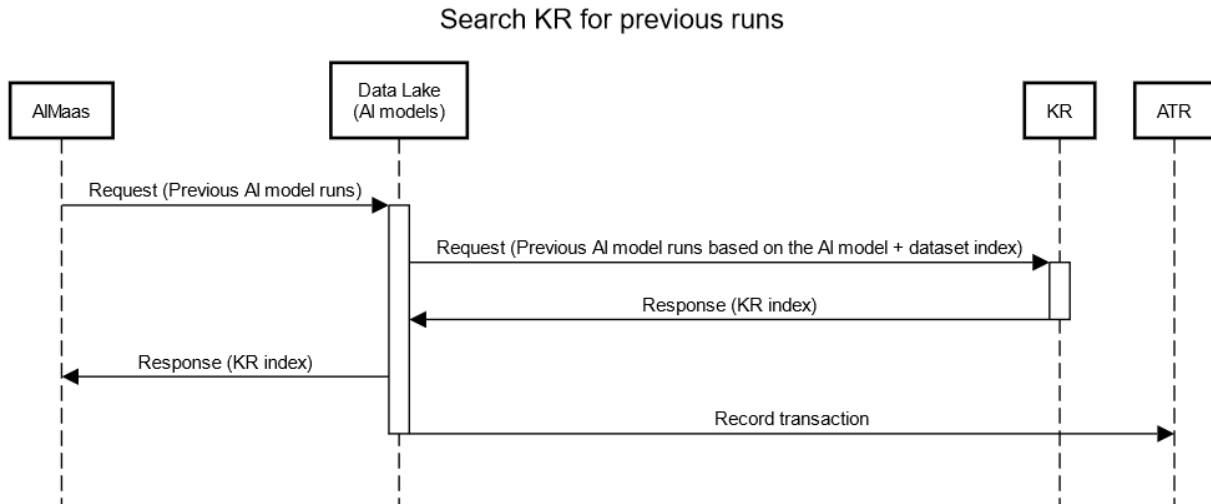


Figure 130: Sequence diagram: DL – KR (search for previous results)

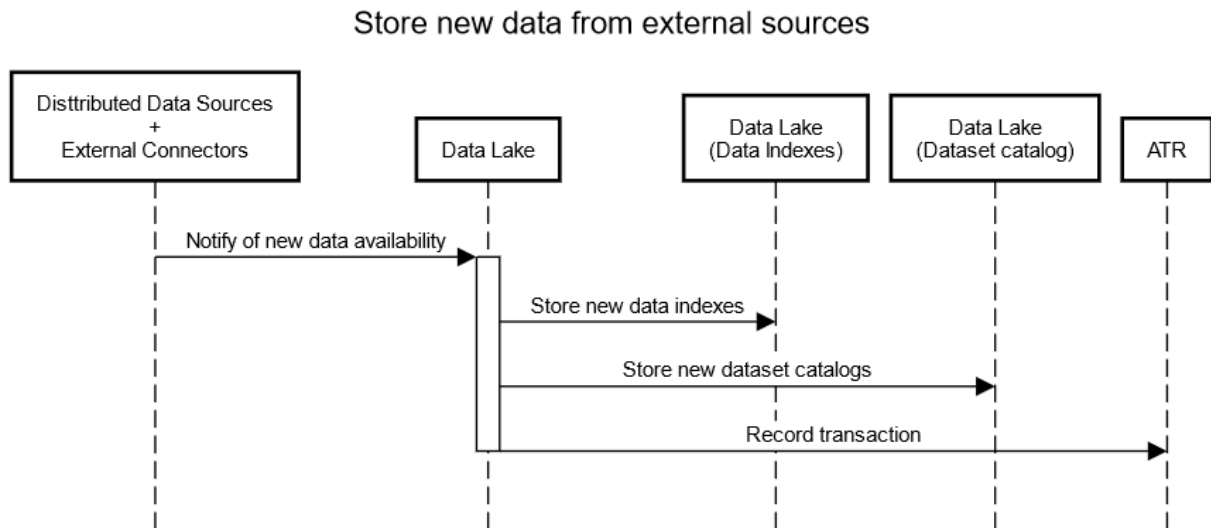


Figure 131: Sequence diagram: DL – Distributed Data Sources & External Connectors (store new data)

ABB External Interfaces specification

Interface	Content	Comment
Interface ID	IExt:DL:HEDF	
Interface to	Homomorphic Enabled Data Fusion - HEDF	
Interface overview	The DL will store the indexes for the datasets that are available from each data provider.	

Messaging pattern	Request/response	
Data exchanged via this interface	Datasets indexes	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DL:SSI-HE-HOLDER	
Interface to	SSI Homomorphic Capable Framework Holder sub-component – SSI-HE-HOLDER	
Interface overview	The DL will provide to SSI-HE ABB the stored homomorphic public keys	
Messaging pattern	Request/response	
Data exchanged via this interface	Public keys	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DL:SSI-HE-ISSUER	
Interface to	SSI Homomorphic Capable Framework Issuer sub-component – SSI-HE-ISSUER	
Interface overview	DL provides the requested homomorphic public keys for metadata	
Messaging pattern	Request/response	
Data exchanged via this interface	Homomorphic public keys for metadata	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DL:SSI-HE-VERIFIER	
Interface to	SSI Homomorphic Capable Framework Verifier sub-component – SSI-HE-VERIFIER	
Interface overview	The DL will provide the datasets' indexes / attributes/features to be searched	
Messaging pattern	Request/response	
Data exchanged via this interface	Endpoints / indexes / attributes	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DL:ATR	
Interface to	Accountable Transactions Recorder – ATR	
Interface overview	The DL will send to the ATR any transaction records carried out.	
Messaging pattern	Request/response	
Data exchanged via this interface	Transactions	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DL:OneSSExternal Connectors	
Interface to	One-Stop-Shop and External Connectors – OneSS External Connectors module	
Interface overview	External data sources' data information for indexing.	
Messaging pattern		To be defined
Data exchanged via this interface		To be defined

Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DLAuthAgent:AM	
Interface to	Authentication Manager – AM	
Interface overview	User authentication information	
Messaging pattern	Request/response	
Data exchanged via this interface	User authentication	
Communication protocol	HTTP(S) / OIDC	
Message format	JSON	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:DL:AIMaaS	
Interface to	AI Models as a Service – AIMaaS	
Interface overview	The DL will provide to the AIMaaS ABB the KR index for any existing model runs for which the KR has information on, and will receive new AI models and logs for storage. AIMaaS will send to DL all trained AI model information (trained model results, datasets indexes used for the training etc.), once the AIMaaS retrieves additional info from the data provider (model description, files necessary for the deployment of the model, the trustworthiness assessment report etc.)	
Messaging pattern	Request/response	
Data exchanged via this interface	AI models, files for the deployment of models, model description, AI model indexes, dataset indexes, logs (datasets used for each AI model's execution)	

Communication protocol		
Message format	json, xml, txt...	
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	

Interface	Content	Comment
Interface ID	IExt:DL:KR	
Interface to	Knowledge Repository – KR	
Interface overview	The DL will search the KR for any previous logs for previous AI models runs (based on the AI model and dataset indexes). If previous run exists, the DL gets the KR index to share to the AIMaaS. KR will send to the DL the index of the last entry run performed by AIMaaS, along with the model and dataset indexes, which will record a new entry in the log with the model, dataset and KR index of the current run.	
Messaging pattern	Request/response	
Data exchanged via this interface	AI model results, KR index, logs, datasets indexes	
Communication protocol		
Message format	json, xml, txt...	
Security and integrity		

ONE-STOP-SHOP (ONESS) BUILDING BLOCK

ABB overview

The TRUSTEE OneSS is envisioned as an online repository (accessible to authorized personnel from TRUSTEE, but also others if needed), to act as the main entry point for the (TRUSTEE) technical people, and eventually the Developer Persona, which refers to a developer outside TRUSTEE with the goal to use exposed functionalities of the TRUSTEE Platform, however, without accessing the data of the Data Providers; that is only use some of the technical solutions developed by the TRUSTEE project. The TRUSTEE OneSS is envisioned to contain the Developer handbook, code repositories, a toolset including code samples, SDKs, libraries, API documentation, and the SDK connectors module.

SDK connectors simplify integration process by providing a unified interface that abstracts the complexity of the underlying technology, in our case cloud-based TRUSTEE external or TRUSTEE internal technology. This enables developers to focus on the functionality of the application rather than the integration details. SDK connectors can

be considered an umbrella term for custom connectors realizing interoperability and integration. Main idea is to achieve efficient interoperability with great user experience for the technical personnel doing the integration.

- SDK external connectors focus on establishing a connection with external system, and the name of the system is used also for the name of the connector (e.g., IDS connector for International Data Space related central technical component for secure and trusted data exchange). As a component that encapsulates access to an external system or resource it can also be considered a *gateway*, as a term typically used by software architects to describe an architecture pattern, denoting a *wrapper* translating the interaction between two systems. From Gang of Four¹⁰⁰⁴ closes pattern is “adapter” but there it is meant in context where both interfaces are already available and here in our case, we are defining our SDK connector interfaces while we wrap external existing one.
- SDK internal connectors enable communication between different components or services within the same system, in our case they are facilitating the integration within TRUSTEE system.
 - One way of achieving this is by utilising OpenAPI specifications. The OpenAPI Specification (currently in v3.1.3 in Feb 2023¹⁰⁰⁵) is a specification for machine-readable interface files for describing, producing, consuming, and visualizing RESTful web services. It was formerly (2009-2017) known as the Swagger Specification.

Some benefits of using OpenAPI include:

- Language-agnostic interface for describing RESTful APIs,
- Machine-readable and easily interpretable by humans,
- Design first approach: whole API with types and examples for every endpoint can be defined before the start of implementation work. With additional tools, mock-APIs can be generated for testing until full implementation becomes available, thus speeding up the work and making agreements explicit
- Code generators: most languages are supported by their code-generators
- Tooling: big ecosystem of tools created under the Swagger brand that can be used to streamline the work (generating API documentation, tests, mock servers), Swagger refers to a set of SmartBear tools (more info on <https://swagger.io/>) which are among the most popular ones for developing and describing RESTful APIs. Swagger tools help users build, document, test and consume RESTful web services. It can be used with both a top-down and bottom-up API development approach.
- Widely accepted: many developers and organization use it,
- Stable: known under Swagger name until 2017. OpenAPI Initiative is part of the Linux Foundation, which increases trustworthiness,
- Reliable single source of truth for APIs: clients and API developers/providers have a good point of reference through the API contract which OpenAPI describes.

ABB application view diagram

The following figure (Figure 132) shows a high-level view of the OneSS building block, focusing on the establishment of the connections with the systems outside TRUSTEE. EOSC and EGI (as well as Copernicus) connectors and related TRUSTEE Use cases were considered but no final decision was made at the time of finalizing

¹⁰⁰⁴ Gamma, Erich, Helm, Richard, Johnson, Ralph and Vlissides, John M.. *Design Patterns: Elements of Reusable Object-Oriented Software*. 1 : Addison-Wesley Professional, 1994.

¹⁰⁰⁵ <https://spec.openapis.org/oas/latest.html>

this deliverable so more work will be invested in this direction in upcoming months of the project and will be further discussed in future versions of the deliverable, namely D2.2 and D2.3, as well as in upcoming deliverables of WP4. Further investigation will also include the legal, organization as well as technical aspects since the TRUSTEE connectors would need connection with a running external system for which several prerequisites have to be achieved.

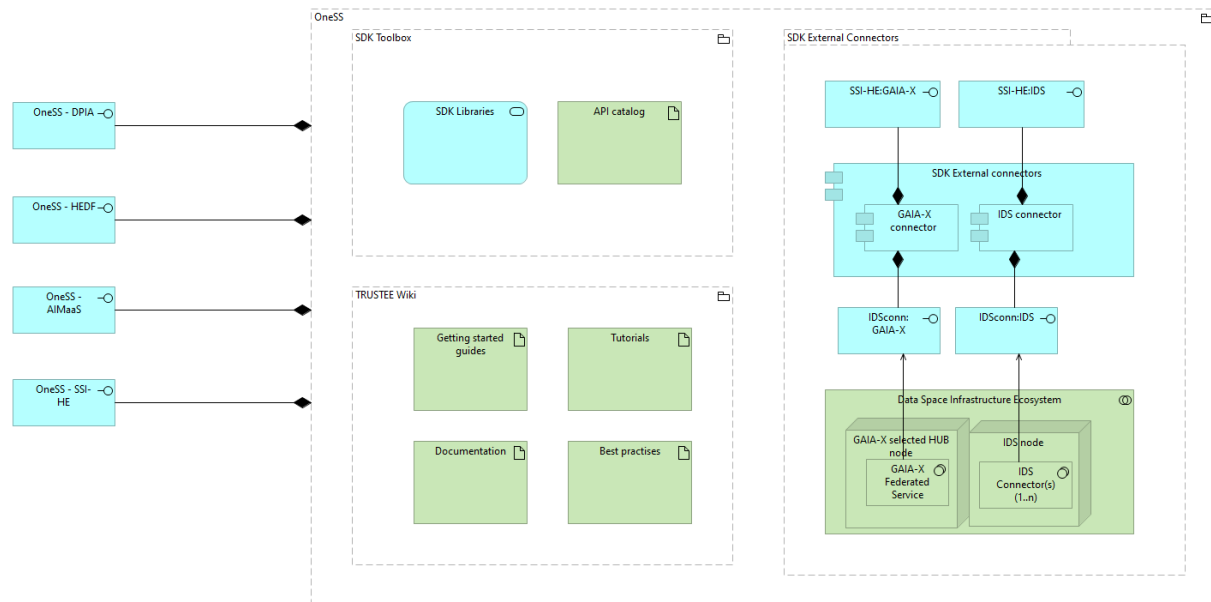


Figure 132: SDK external connectors ABB breakdown

ABB Sequence Diagrams

As this ABB requires further investigation as indicated above, the sequence diagrams pertaining to the external communication of this ABB with other TRUSTEE ABBs as well as with external connectors through the use of the SDK External Connectors inner module will be included in the next version of this deliverable, namely D2.2 and will be updated if needed in D2.3. Additional details about this ABB will be provided in D4.1.

ABB External interfaces specification

Interface	Content	Comment
Interface ID	IExt:IDScnncor	
Interface to		Initially not connected to other entities. Connections can be dynamically established with other IDS connectors.
Interface overview	Connector is the technical component for secure and trusted data exchange and it provides the technical access to the IDS ecosystem. The connector uses application container	

	management technology (docker container) to provide a secure and isolated environment for individual data services. In IDS data is always exchanged from one connector to another in point-to-point manner. The connector enables data providers to define rules and conditions (usage policies) under which data is shared with a data consumer.	
Messaging pattern	Request/response, publish-subscribe, notification	
Data exchanged via this interface	The connector offers three different types of data types for other connectors: - Arbitrary strings: can be, for example, JSON statements. - Files: files can be selected from the host machine file system. - Remote sources: data can be dynamically fetched from remote REST APIs. Additionally, data can be extracted from OPC UA servers. There are no restrictions for the structure of the data that is exchanged (i.e., the “payload”). However, the metadata of every message needs to be aligned with the IDS information model.	
Communication protocol	IDS Communication Protocol (IDSCP), HTTP Multipart	
Message format	Not restricted (see above section “Data exchanged via this interface”)	
Security and integrity	The connector implements the IDS reference architecture security (e.g., identity and access management) mechanisms. Every connector must have an IDS certificate, which is based on the X509v3 standard. Connectors authenticate themselves using the certificates in the IDS Identity Provider module. The Identity Provider then issues OAuth2 access tokens for the connectors, thus enabling the connectors to access the services and data of other connectors. Data provider connectors can restrict access to their offered data resources on a connector and a user level. This means that a data provider can either define i) consumer connectors or ii) specific users of data consumer connectors that are allowed to access specified data resources.	

Interface	Content	Comment

Interface ID	<i>IExt:EOSCconnector</i>	Still under consideration, no final decision yet
Interface to	<i>European Open Science Cloud (EOSC)</i>	
Interface overview	EOSC is a federated and open multi-disciplinary environment where they can publish, find and re-use data, tools and services for research, innovation and educational purposes	
Messaging pattern		
Data exchanged via this interface		
Communication protocol		
Message format		
Security and integrity		

Interface	Content	Comment
Interface ID	<i>IExt:EGIconnector</i>	Still under consideration, no final decision yet
Interface to	<i>European Grid Infrastructure (EGI)</i>	
Interface overview	<i>Through the connector it's possible to mount a space locally, and access it over a POSIX interface, using files as they were stored locally. The file's blocks are downloaded on demand.</i>	
Messaging pattern	<i>Request/Response</i>	
Data exchanged via this interface		
Communication protocol	<i>RESTful (HTTPS)</i>	
Message format	<i>JSON</i>	
Security and integrity	<i>The EGI implements an Authentication & Authorization procedure (EGI Check-In) which is based on the AARC Blueprint Architecture.</i>	

HOMOMORPHIC ENABLED DATA FUSION (HEDF) BUILDING BLOCK

ABB overview

The HEDF Building Block will provide a space for communicating with the SSI_HE ABB and will include the HE Toolbox, which will comprise 3 sub-components (HE Manager, HE Module and the Secure Enclave).

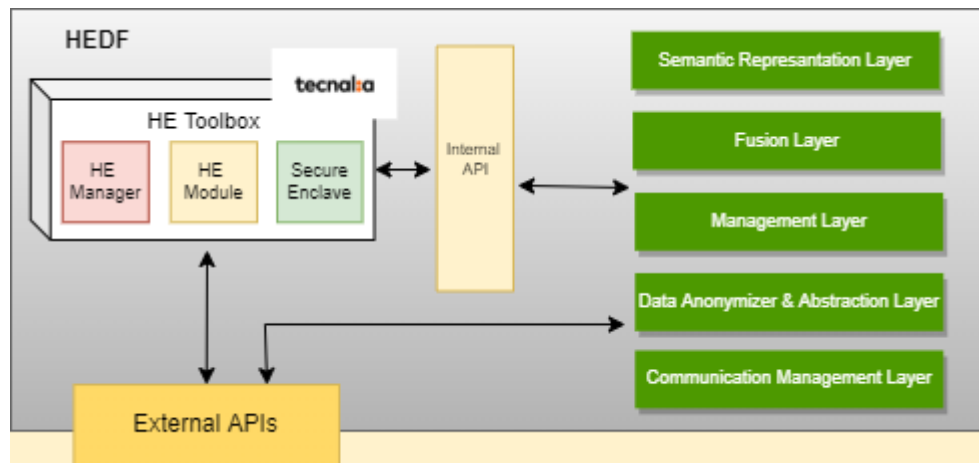


Figure 133: HEDF ABB architecture

The HEDF middleware will consist of 4 layers:

- Semantic representation layer, which will be responsible for the transformation of the collected data and models into knowledge, by applying common/standardized semantic ontologies on the data, correlated with the appropriate contextual information.
- Fusion layer, which will be responsible for the data fusion process, inter-connecting with the HE Toolbox.
- Management Layer, which will manage the incoming data events, and categorize the data appropriately, to allow for better processing and comprehension from the TRUSTEE Platform.
- Communication Management Layer, which will provide different communication drivers, to allow gathering of data over different interfaces and protocols.

The HEDF ABB will communicate with the following external interfaces and blocks:

- SSI-HE Verifier (IExt:HEDF:SSI-HE-VERIFIER). HEDF will receive from the SSI-HE VERIFIER the data providers' endpoints / indexes that hold the data in order to start the fusion process
- SDK (IExt:HEDF:SDK). HEDF receives from SDK external connectors for accessing data from external sources (ISD, EOSC, EGI connectors).
- AIMaaS (IExt:HEDF:AIMaaS). HEDF will run the AIMaaS models. When an AI model is trained with HE data within HEDF, the trained model, e.g. a file, will be sent back to AIMaaS to be stored, along with the dataset indexes that were used for training.
- DA (IExt:HEDF:DA). HEDF will provide the DA with the results of the computation that the TRUSTEE User requested.
- DL (IExt:HEDF:DL). HEDF will receive from the TRUSTEE DL the stored datasets indexes / AI models
- ATR (IExt:HEDF:ATR). HEDF will send to ATR for recording all performed activities
- KR (IExt:HEDF:KR). HEDF will check if KR has stored results from data processing / computations from prior sessions in order to reduce processing load.
- STM (IExt:HEDF:STM). HEDF, upon a data consumer's request, checks from STM agreements on datasets, agreements on models, privacy policies and DPIA regarding a specific computation.

ABB application view diagram

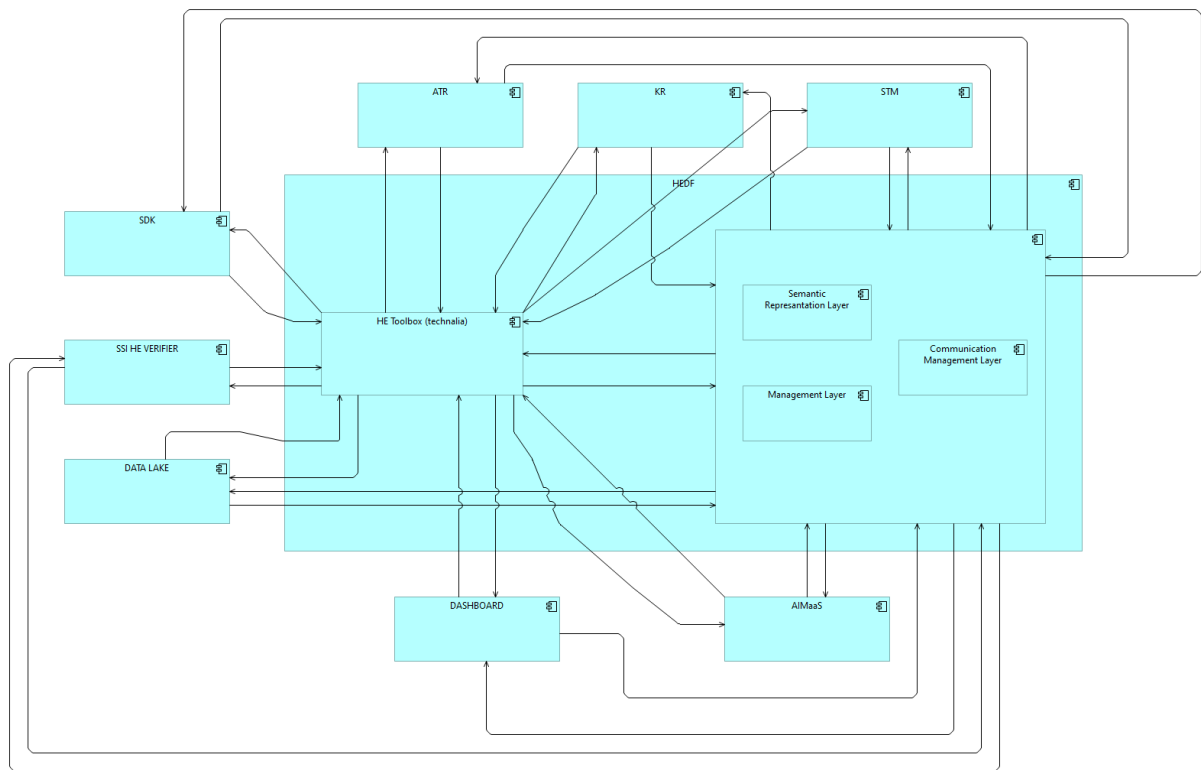


Figure 134: HEDF ABB diagram

ABB Sequence Diagrams

The sequence diagram depicted in Figure 135 presents an initial version of the communication of HEDF with the various ABBs of the TRUSTEE Platform for performing computation results over the encrypted data, realizing the request coming from a Consumer.

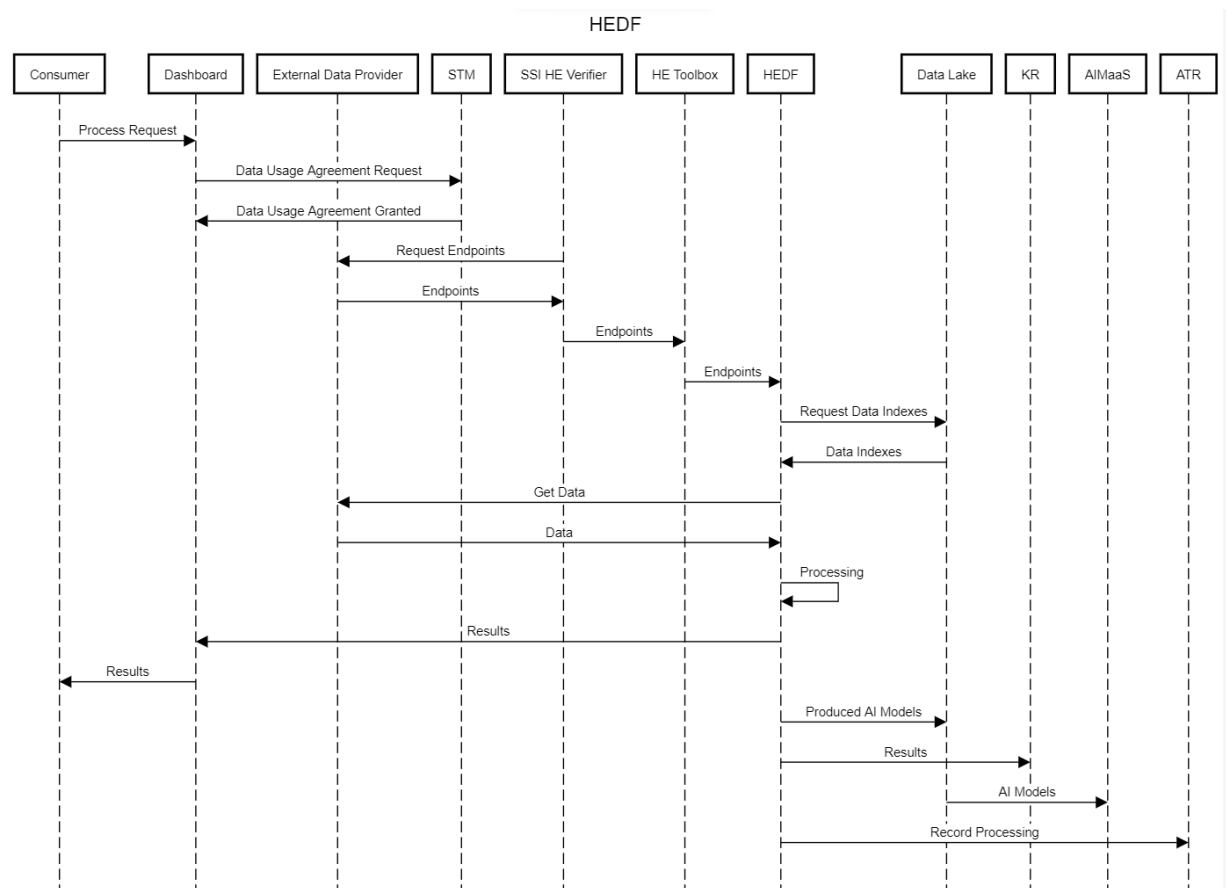


Figure 135: HEDF Sequence diagram

ABB External interfaces specification

Interface	Content	Comment
Interface ID	IExt:HEDF:DA	
Interface to	Dashboard – DA	
Interface overview	The DArequests from HEDF the results of the computation that the TRUSTEE User requested	
Messaging pattern	Request/Response	
Data exchanged via this interface	HE data & Data Fusion results	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:HEDF:SDK	
Interface to	One Stop Shop External Connectors module – OneSSEExternalConnectors	
Interface overview	The external connectors module provides HEDF with external connectors for accessing data from external sources (ISD, EOSC, EGI connectors)	
Messaging pattern	Request/Response	
Data exchanged via this interface	Depending on the External Connector's features (IExt:IDSconnector, IExt:EGIconnector, IExt:EOSCconnector)	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure communications (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:HEDF:STM	
Interface to	Security and Trust Manager (Agreement management platform) – STM	
Interface overview	HEDF, upon a data consumer's request, checks from STM agreements on datasets, agreements on models, privacy policies and DPIA regarding a specific computation.	
Messaging pattern	Request/response	
Data exchanged via this interface	Meta-data	
Communication protocol	Free to secure the communications	
Message format	JSON	
Security and integrity	Trustee platform	

Interface	Content	Comment
Interface ID	IExt:HEDF:SSI-HE-VERIFIER	

Interface to	SSI Homomorphic Capable Framework Verifier module – SSI-HE-VERIFIER	
Interface overview	HEDF will receive from the SSI-HE VERIFIER the data providers' endpoints / indexes that hold the data in order to start the fusion process	
Messaging pattern	GET Request	
Data exchanged via this interface	Endpoints & indexes	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure communication (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:HEDF:DL	
Interface to	Data Lake – DL	
Interface overview	HEDF will receive from the TRUSTEE DL the stored datasets indexes / AI models	
Messaging pattern	GET Request	
Data exchanged via this interface	Endpoints & indexes	
Communication protocol	HTTPS	
Message format	JSON	
Security and integrity	Secure communication (HTTPS)	

Interface	Content	Comment
Interface ID	IExt:HEDF:KR	
Interface to	Knowledge Repository – KR	
Interface overview	HEDF will check if KR has stored results from data processing / computations from prior sessions in order to reduce processing load	
Messaging pattern	Request / Response	
Data exchanged via this interface	Results stored in KR (computational results, datasets / AI models indexes, timestamps)	

Communication protocol	HTTP, REST	
Message format	JSON, XML, YAML	
Security and integrity		

Interface	Content	Comment
Interface ID	IExt:HEDF:AIMaaS	
Interface to	AI Models as a Service – AIMaaS	
Interface overview	HEDF will run the AIMaaS models. When an AI model is trained with HE data within HEDF, the trained model, e.g. a file, will be sent back to AIMaaS to be stored, along with the dataset indexes that were used for training	
Messaging pattern	Request/response	
Data exchanged via this interface	Data index, model index, trained model, training results, inference results	
Communication protocol		
Message format		
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	

Interface	Content	Comment
Interface ID	IExt:HEDF:ATR	
Interface to	Accountable Transactions Recorder – ATR	
Interface overview	HEDF will send to ATR for recording all performed activities	
Messaging pattern	Request/response	
Data exchanged via this interface	Datasets / AI models for processing, HE fusion	
Communication protocol		
Message format		
Security and integrity	local infrastructure (private cloud), or Trustee SDK, or Trustee platform	

KNOWLEDGE REPOSITORY (KR) BUILDING BLOCK

ABB overview

The TRUSTEE Knowledge Repository (KR) will:

- be implemented utilising InterPlanetary File System (IPFS) and will have the ability to store knowledge-related data stemming from search queries, results of computations over the encrypted data of Data Providers, and information about AI/ML use and/or inference run results (if the Data Providers and Models Providers have agreed on this).
- be established within a distributed peer-to-peer private network to facilitate the storage and sharing of content in a distributed file system.
- be based on a content addressing approach where every stored content will generate a unique hash which will work as an identifier of the content.
- comprise an automated garbage collection functionality which will enable permanent removal of unused or orphaned contents from the repository.
- provide a pinning option to ensure the availability and preservation of historical knowledge.
- be integrated based on data homogenization and semantic alignment toolkits that will be researched and developed.
- be integrated with external tools for enhanced functionalities such as knowledge graphs, search capability, etc.
- comprise automated analysis functionalities which should provide the ability to assist users with the query. For instance, the queries that are similar to the previous queries made. In addition, the main idea of the KR is reducing energy consumption. The result exists from the previous query, and they are voluntarily put inside the KR.
- be able to replicate and provide the results to the end users (such as the Consumer).

ABB application view diagram

Figure 136 depicts a distributed architecture comprising both IPFS peer-to-peer protocol and a blockchain network. In this architecture, an entity uploads knowledge-related content to the IPFS network and receives the content identifier (CID). The distributed hash table (DHT) provides a look-up service of content stored in peer nodes. The CID is subsequently recorded to the blockchain, being appended to a new block. The stored content can be retrieved using the CID.

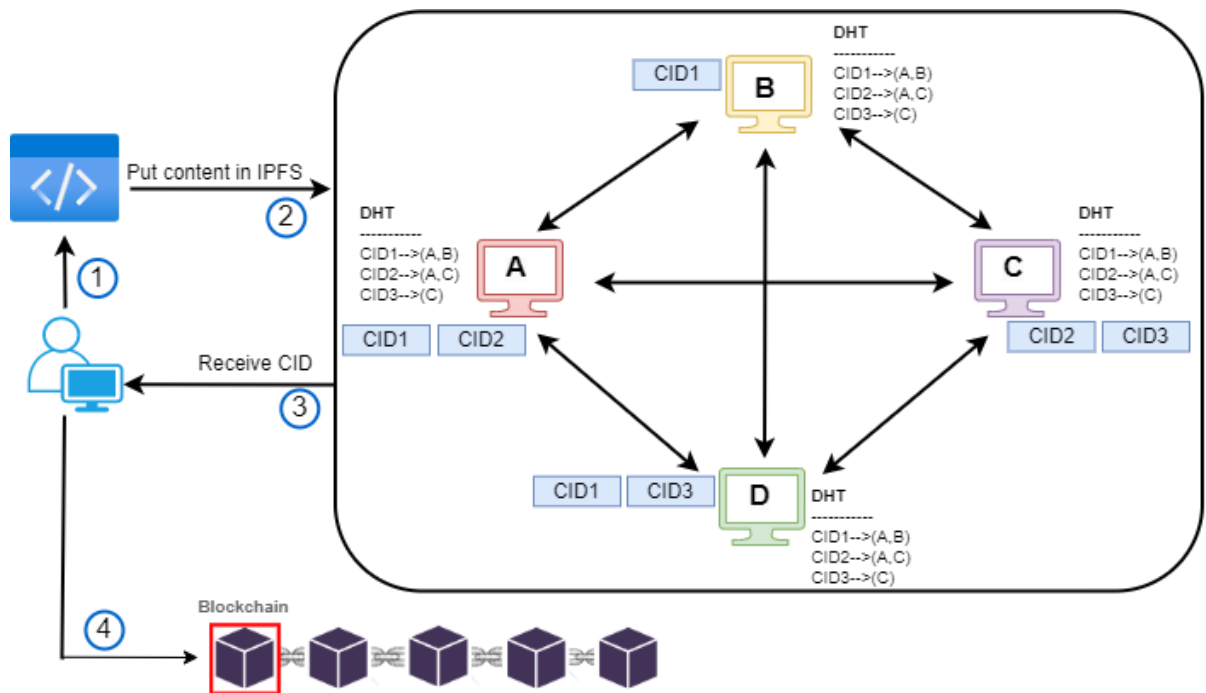


Figure 136: Content gets stored to an IPFS node and the generated CID is sent to the USER. The CID is then stored in the blockchain.

Following, Figure 137 showcases the specification of Knowledge repository and external interfaces with the various ABBs of the TRUSTEE Platform.

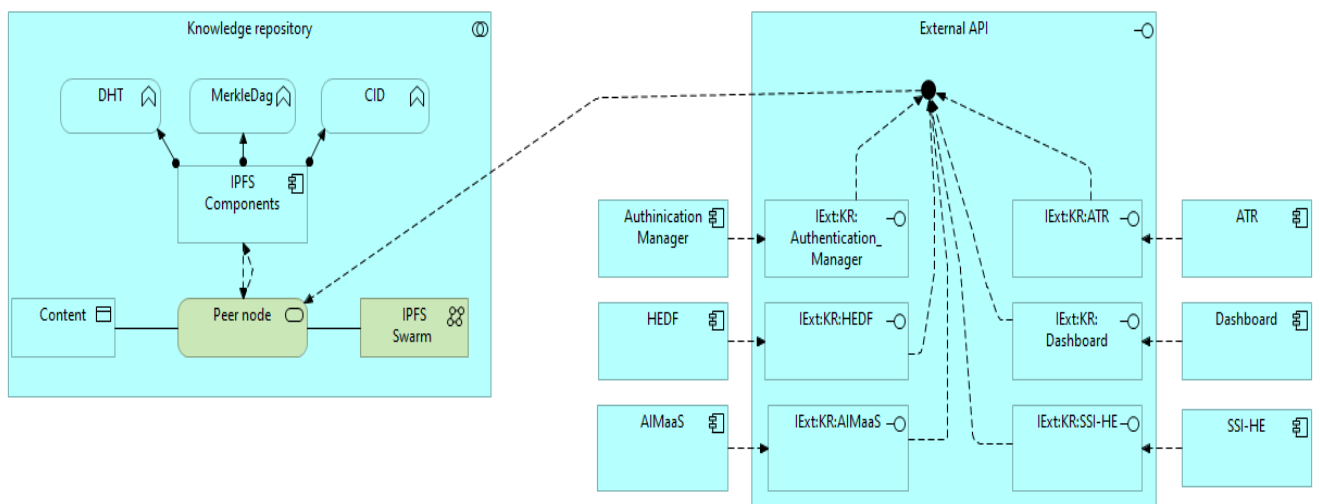


Figure 137: Knowledge Repository decomposition diagram and external interfaces

ABB Sequence Diagrams

Figure 138 below shows knowledge related content stored in KR, while the generated CID is recorded to the ATR. The recorded transaction can be obtained from the Dashboard. The steps to access the stored contents from the KR is depicted in a sequence diagram in Figure 138.

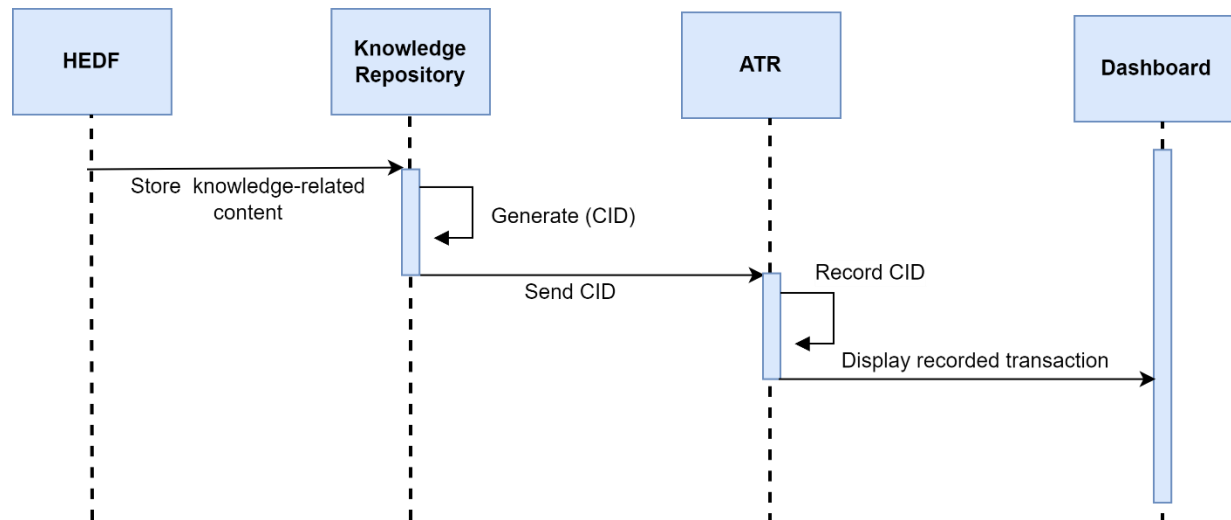


Figure 138: Content stored at Knowledge Repository and ATR integration for transaction recording

Additionally, the sequence diagram illustrating the process of a user requesting to view stored knowledge-related content from the KR through the DA has been included in the TRUSTEE Dashboard (DA) Building Block section.

ABB External interfaces specification

Interface	Content	Comment
Interface ID	<i>IExT:KR:ATR</i>	The ATR will be integrated with the Knowledge Repo (IPFS)
Interface to	Accountable Transaction Recorder – ATR	
Interface overview	The ATR will record the hash (CID) of every knowledge-related content that is uploaded to the KR.	
Messaging pattern	Request/Response	
Data exchanged via this interface	Timestamps, CID (Content Identifier)	
Communication protocol	gRPC, HTTP, REST, Kafka, IPFS API	

Message format	JSON, protobuf, CBOR, XML, YAML	
Security and integrity	Decentralised infrastructure. No central authority is in charge; thus no single point of failure.	

Interface	Content	Comment
Interface ID	<i>IExT:KR:DA</i>	
Interface to	Dashboard – DA	
Interface overview	The user (Consumer) will be able to view knowledge extracted from past queries or use of AI/ML models by them or by other users	
Messaging pattern	Request/Response	
Data exchanged via this interface	Knowledge-related content (e.g., results of past queries, results of past computations over the encrypted data)	
Communication protocol	HTTP(S)	
Message format	JSON	
Security and integrity	Decentralised infrastructure. No central authority is in charge; thus, no single point of failure.	

Interface	Content	Comment
Interface ID	<i>IExT:KR:HEDF</i>	
Interface to	Homomorphic Enbaled Data Fusion – HEDF AI Model as a Service – AIMaaS SSI Homomorphic Capable Framework Verifier module – SSI-HE-VERIFIER	Many of the TRUSTEE ABBs will need to use this interface for knowledge-related content storing.
Interface overview	The goal is to store content and obtain CID.	
Messaging pattern	Request/Response	
Data exchanged via this interface	Timestamps, CID (Content Identifier)	
Communication protocol	gRPC, HTTP, REST, Kafka, IPFS API	
Message format	JSON, protobuf, CBOR, XML, YAML	

Security and integrity	Decentralised infrastructure. No central authority is in charge; thus no single point of failure.	

TECHNOLOGY, PHYSICAL VIEWS

In the ArchiMate modelling language, the Technology and Physical views provide a comprehensive picture of the technology infrastructure of an enterprise (in our case TRUSTEE), from the high-level components to the physical deployment details (see chapter ArchiMate Full Framework). This information can be used to understand the impact of changes to the technology infrastructure, to pinpoint areas for improvement and optimization, and to plan for future technology investments.

- The Technology view is used to describe the technology infrastructure of an enterprise (in our case TRUSTEE), including the hardware, software, and network components. It shows the technology components and their relationships and dependencies, as well as their interactions with the business and application components.
- The Physical view is used to describe the physical deployment of the technology infrastructure, including the hardware and network topology, the locations of servers, and the physical connections between components. This view is useful for understanding how the technology infrastructure is implemented and how it relates to the physical environment.

INFRASTRUCTURE VIEW

The Infrastructure View in ArchiMate is a view that focuses on the physical and technical aspects of an enterprise's (in our case TRUSTEE) infrastructure. It provides a holistic overview of the hardware, software, network, and communication components.

The TRUSTEE Infrastructure View is a critical view for TRUSTEE IT architects and technical stakeholders to ensure that the TRUSTEE infrastructure is designed, built, and maintained to support the TRUSTEE operations effectively. It provides a comprehensive and detailed understanding of the enterprise's infrastructure and its interdependencies, which can help to identify potential issues, dependencies, and risks.

Following figure (Figure 139) shows the infrastructure serving the needs of the TRUSTEE Applications, belonging to three layers, namely:

- The application layer, containing TRUSTEE applications, represents the software and applications that run on the TRUSTEE infrastructure.
- The physical layer represents the physical components of the infrastructure, such as servers, storage devices, and other hardware components.
- The communication layer represents the network and communication infrastructure, including routers, switches, and other network devices.

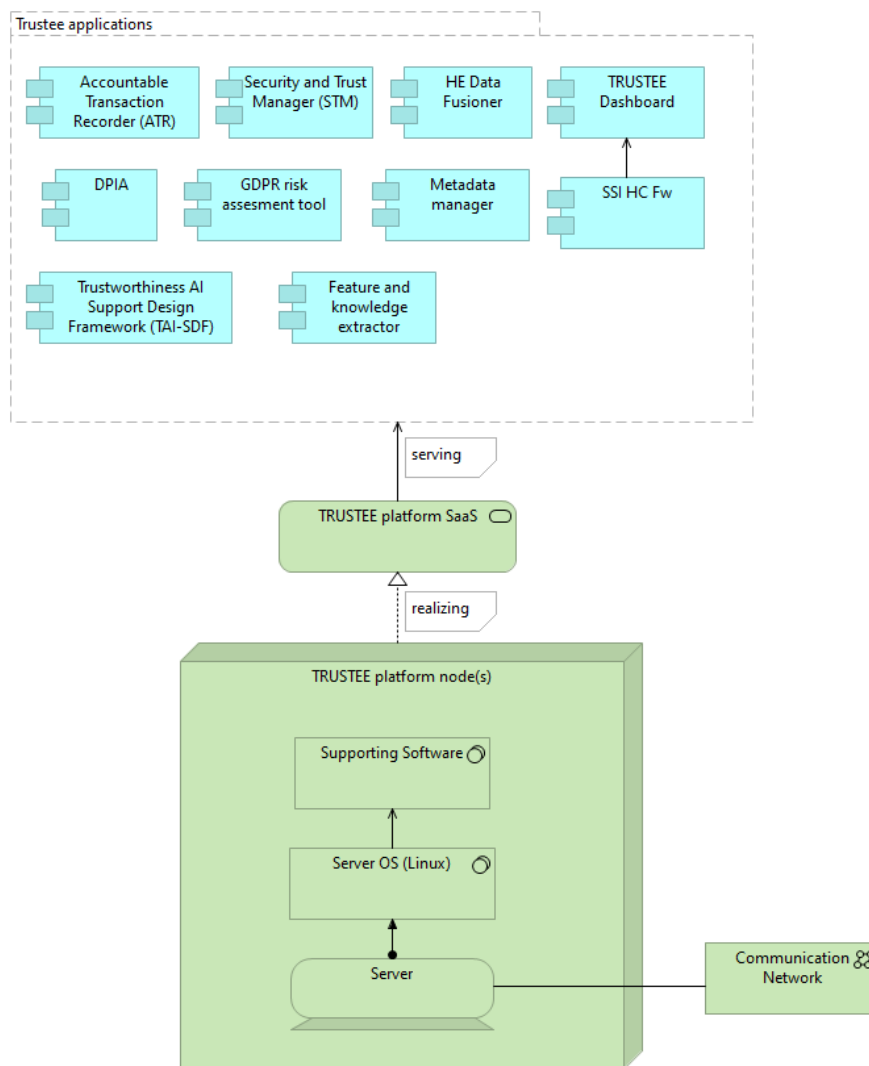


Figure 139: TRUSTEE infrastructure serving the needs of the selected TRUSTEE Applications

The technology view will be further elaborated and detailed within the implementation WPs (WP3, WP4) as it represents the next step towards the implementation and deployment of the TRUSTEE functionalities.

IMPLEMENTATION VIEWS: FEDERATED LEARNING CASE

Implementation views are primarily addressed in the WP3 and WP4, however for completeness of the selected ArchiMate Full Framework we explore in this deliverable the FL case, which is also the focus of T2.5 within WP2. The FL case is investigated for the automotive data domain.

In the following, we present:

- A block diagram of the envisioned ATHENA's components with their connections and task assignments.
- For each component, we provide a short description and the requested table with the external interface specifications.

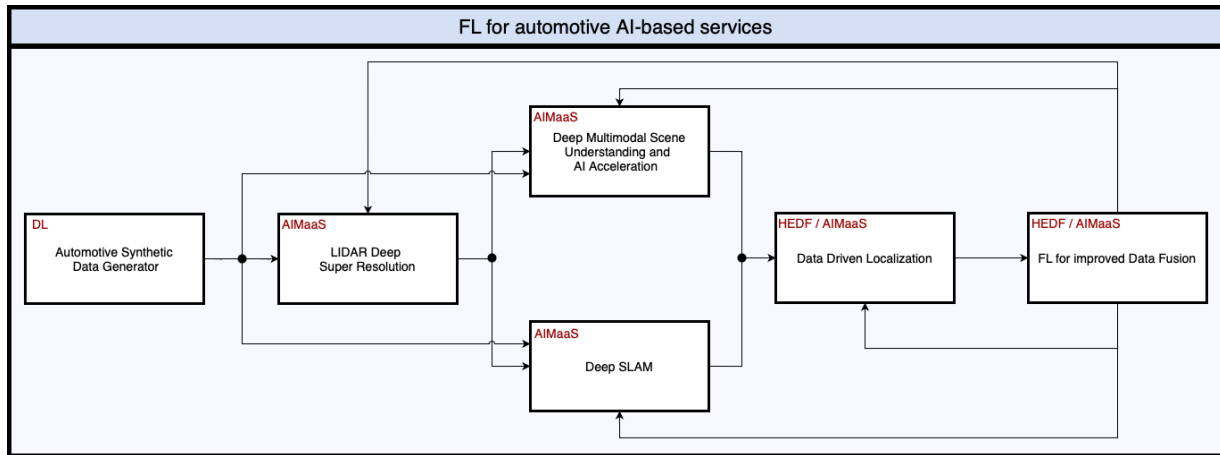


Figure 140: FL case for automotive data domain: ABB components and connections

Components

DEEP MULTIMODAL SCENE ANALYSIS

This module is responsible for detecting and identifying objects of interest in automotive scenes, using convolutional 2D and 3D DNNs. Sensors like cameras and LIDARs mounted on simulated vehicles will feed those models with 2D images and 3D point clouds. Additionally, to ease the burden of computational complexity, advanced AI acceleration techniques will be employed in this module in order to compress and accelerate the involved scene analysis and understanding DNNs. Finally, utilizing geometric properties, late fusion of camera's and LIDAR's bounding boxes will be applied so as to improve the scene analysis ability.

Interface	Content	Comment
Interface ID	DMSA	
Interface to	ASDG, LDSR, DDL, FLAAS	
Interface overview	This module is responsible for detecting and identifying objects of interest in automotive scenes, using convolutional 2D and 3D DNNs.	
Messaging pattern	publish-subscribe	
Data exchanged via this interface	Input: Image data (png, jpg), Point cloud data (bin, ply), ROS messages (sensor_msgs/Image, sensor_msgs/PointCloud2), rosbag Output: txt file (KITTI format), ROS messages	
Communication protocol	Based on the simulation setup due to HW in the loop, DDS (ROS2)	
Message format	ROS messages	

Security and integrity	ROS2 based security specifications for enclave identity, enclave permission and private keys	

DEEP SLAM

This module will be responsible for estimating the pose of the ego vehicle and simultaneously creating a map of the environment. It will diversify from traditional SLAM systems since it will employ deep learning-based components. In the case of visual Deep SLAM the component will get as input camera images from ASDG and its output will be available to DDL. Similarly, in the case of LiDAR Deep Slam the input would be a point cloud directly from ASDG or its upscaled version from LDSR. In the context of FL, the system will participate in a decentralized training scheme in which it will exchange machine learning models with FLAAS.

Interface	Content	Comment
Interface ID	DS	
Interface to	DDL, LDSR, ASDG, FLAAS	
Interface overview	This module will be responsible for estimating the pose of the ego vehicle and simultaneously creating a map of the environment.	
Messaging pattern	publish-subscribe	
Data exchanged via this interface	Input: Image data (png, jpg), ROS messages (sensor_msgs/Image), Rosbag, deep learning model (pth) Output: txt file (KITTI format), csv file ROS messages (sensor_msgs/PointCloud2, geometry_msgs/PoseStamped), deep learning model (pth)	
Communication protocol	Based on the simulation setup due to HW in the loop, DDS (ROS2)	
Message format	ROS, txt, csv	
Security and integrity	ROS2 based security specifications for enclave identity, enclave permission and private keys (link), extra level FL, extra level HE	

DATA DRIVEN LOCALIZATION

This module will be responsible for enabling cooperative location awareness over a swarm of connected and autonomous vehicles, both in a distributed and centralized manner. The two previous modules of Deep Multimodal

Scene Analysis and SLAM will assist this one, so as to the relative and self-measurement models among vehicles to encode the location estimation scheme. Additionally, pre-trained models will be key-operational parts of this module to exploit the uncertainty of measurements in the data drive localization algorithm.

Interface	Content	Comment
Interface ID	DDL	
Interface to	ASDG, DMSA, DS, LDSR	
Interface overview	This module will be responsible for enabling cooperative location awareness over a swarm of connected and autonomous vehicles, both in a distributed and centralized manner.	
Messaging pattern	publish-subscribe	
Data exchanged via this interface	Input: output from DMSA, LDSR, rosbag Output: txt file (for each vehicle of the swarm: id, 3D position, 3D rotation angle)	
Communication protocol	Based on the simulation setup due to HW in the loop, DDS (ROS2)	
Message format	ROS messages	
Security and integrity	ROS2 based security specifications for enclave identity, enclave permission and private keys (link)	

LIDAR DEEP SR

This module will be responsible for reconstructing high-resolution 3D point clouds utilizing only as input a low-resolution lidar sensor. In particular, having only a low-resolution lidar with 16 channels instead of using sparse point clouds, we can produce more dense point clouds based on highly interpretable and concise deep learning architectures derived from the deep unrolling strategy. Then, the estimated dense point cloud can be used as input to other algorithms for SLAM and scene analysis to increase their performance. Thus, this module will assist the Deep Multimodal and SLAM modules providing point clouds that are able to capture more details about the environment.

Interface	Content	Comment
Interface ID	LDSR	
Interface to	DMSA, DS	
Interface overview	This module will be responsible for reconstructing high-resolution 3D point clouds utilizing only as input a low-resolution lidar sensor.	

Messaging pattern		
Data exchanged via this interface	Input: output from ASDG, Point cloud data (bin, ply), ROS messages (sensor_msgs/PointCloud2), Rosbag Output: Point cloud data (bin, ply), deep learning model (pth)	
Communication protocol	Based on the simulation setup due to HW in the loop, DDS (ROS2)	
Message format		
Security and integrity	Ros, extra level FL, extra level HE	

FL FOR AUTOMOTIVE AI BASED SERVICES

This module will be responsible for enabling spatially dispersed edge devices to train collaboratively a deep learning global model without the devices sharing any sensitive information regarding their local datasets, in a manner which is orchestrated by the centralized server. Under the FL framework, we will examine various problems related with the lidar super-resolution, scene analysis, deep slam and data driven localization. Hence, this module will assist the Lidar Deep SR, Deep Multimodal, Deep SLAM and Data Driven Localization modules. Additionally, the HE method will be employed to decrypt the weights of the local models of the devices for an additional level of security. Finally, this module will enable advanced data fusion techniques based on FL, for improving the major tasks of the previous modules. This task is realized by the submodule of FL for improved data fusion shown in the architecture diagram of the proposed simulated framework.

Interface	Content	Comment
Interface ID	FLAAS	
Interface to	DMSA, DS, LDSR, DDL	
Interface overview	This module will be responsible for enabling spatially dispersed edge devices to train collaboratively a deep learning global model without the devices sharing any sensitive information regarding their local datasets	
Messaging pattern		
Data exchanged via this interface	Input: output from DMSA, DS, DDL, LDSR (deep learning models (pth)) Output: Deep learning model (pth)	
Communication protocol	Based on the simulation setup due to HW in the loop, DDS (ROS2)	
Message format		

Security and integrity	Ros, extra level FL, extra level HE	

AUTOMOTIVE SYNTHETIC DATA GENERATOR

This module will be capable of providing synchronized and annotated data, mainly targeting FL techniques for automotive scenarios. It will provide data that could be produced by different sources like simulated sensors or perception algorithms. The simulator enables us to acquire the ground truth from the system and thus we can obtain a fully annotated dataset. In addition, it will generate diverse data, representative of real-world scenarios, realistically distributed, balanced and adequately sized to ensure the effective training of the machine learning algorithms.

Interface	Content	Comment
Interface ID	ASDG	
Interface to	DMSA, LDSR, DS	
Interface overview	This module will be capable of providing synchronized and annotated data, mainly targeting FL techniques for automotive scenarios	
Messaging pattern	publish-subscribe	
Data exchanged via this interface	Input: automotive scenario descriptions (ASAM OpenSCENARIO format) Output: Point cloud data (bin, ply), Images (png, jpg), ROS messages (sensor_msgs/PointCloud2, sensor_msgs/Image), V2X messages (ETSI ITS-G5) Rosbag, Text files (csv, txt)	
Communication protocol	Based on the simulation setup due to HW in the loop.	
	DDS (ROS2)	
Message format	XML, CSV, ROS	
Security and integrity	ROS, extra level FL, extra level HE	

MOTIVATION VIEWS

Motivation elements are used to model the motivations, or reasons, which guide the design or change of an Architecture.¹⁰⁰⁶

MOTIVATION LAYERS VIEW

The Motivational Layers view in ArchiMate is used to model the drivers, goals, principles, and stakeholders that motivate or guide the development and implementation of architecture. Overall, the Motivational Layers view is a useful tool for understanding the strategic context and drivers of an organization, and for aligning the development and implementation of architecture with the organization's goals and objectives. The TRUSTEE Motivational Layers view shown in following figure focuses on the intangible aspects of TRUSTEE, such as values, and strategic objectives, rather than the tangible elements.

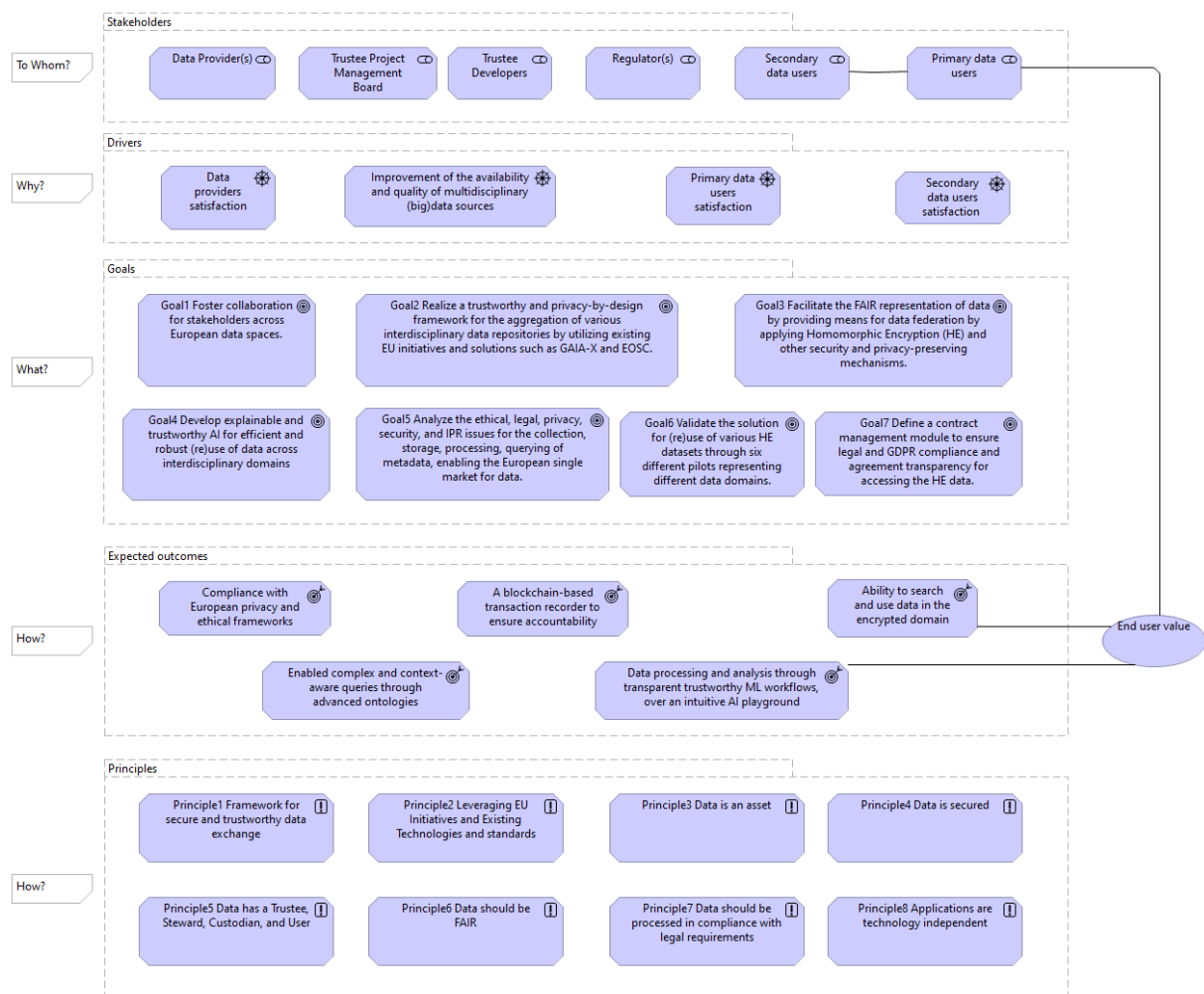


Figure 141: TRUSTEE Motivational layers view

¹⁰⁰⁶ ArchiMate® 3.1 Specification, Copyright © 2012-2019 The Open Group, Motivation Elements <https://pubs.opengroup.org/architecture/archimate3-doc/chap06.html#>

STAKEHOLDERS TO GOALS

By modelling stakeholders, drivers, and goals in the Motivational Layers view, all TRUSTEE partners and members, can better understand the relationships between these elements and ensure that architecture development is aligned with the interests and goals of stakeholders.

Following figure shows TRUSTEE Motivation layers from (internal or external) TRUSTEE Stakeholders (who have an interest in or are affected by the TRUSTEE architecture) to Goals (as the desired outcomes or objectives that the TRUSTEE consortium wants to achieve through the development and implementation of TRUSTEE architecture).

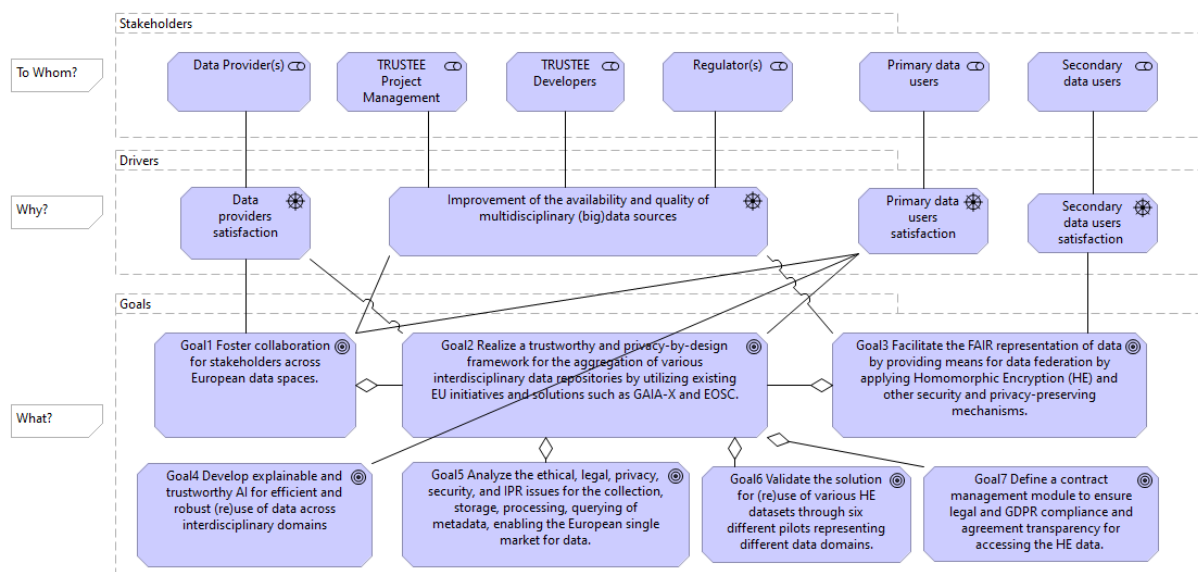


Figure 142: TRUSTEE Motivation layers from Stakeholders to Goals

TRUSTEE HOLISTIC CONCEPTUAL ARCHITECTURE

In addition to the previous ArchiMate compliant views of the architecture in this section we present overall TRUSTEE conceptual architecture view (Figure 143) illustrating high-level design. It does not contain 1:1 mapping between previously described ABBs but instead focuses on providing concise representations of the subset of critical functionalities offered by the TRUSTEE (Federated) Platform.

The idea behind this illustration is to facilitate shared understanding of the TRUSTEE architecture between all stakeholders, to guide the development and evolution of the TRUSTEE Federated Platform over time, to reduce the complexity and risk of system development by providing clear design, blueprint, that is easy to follow, understand and maintain.

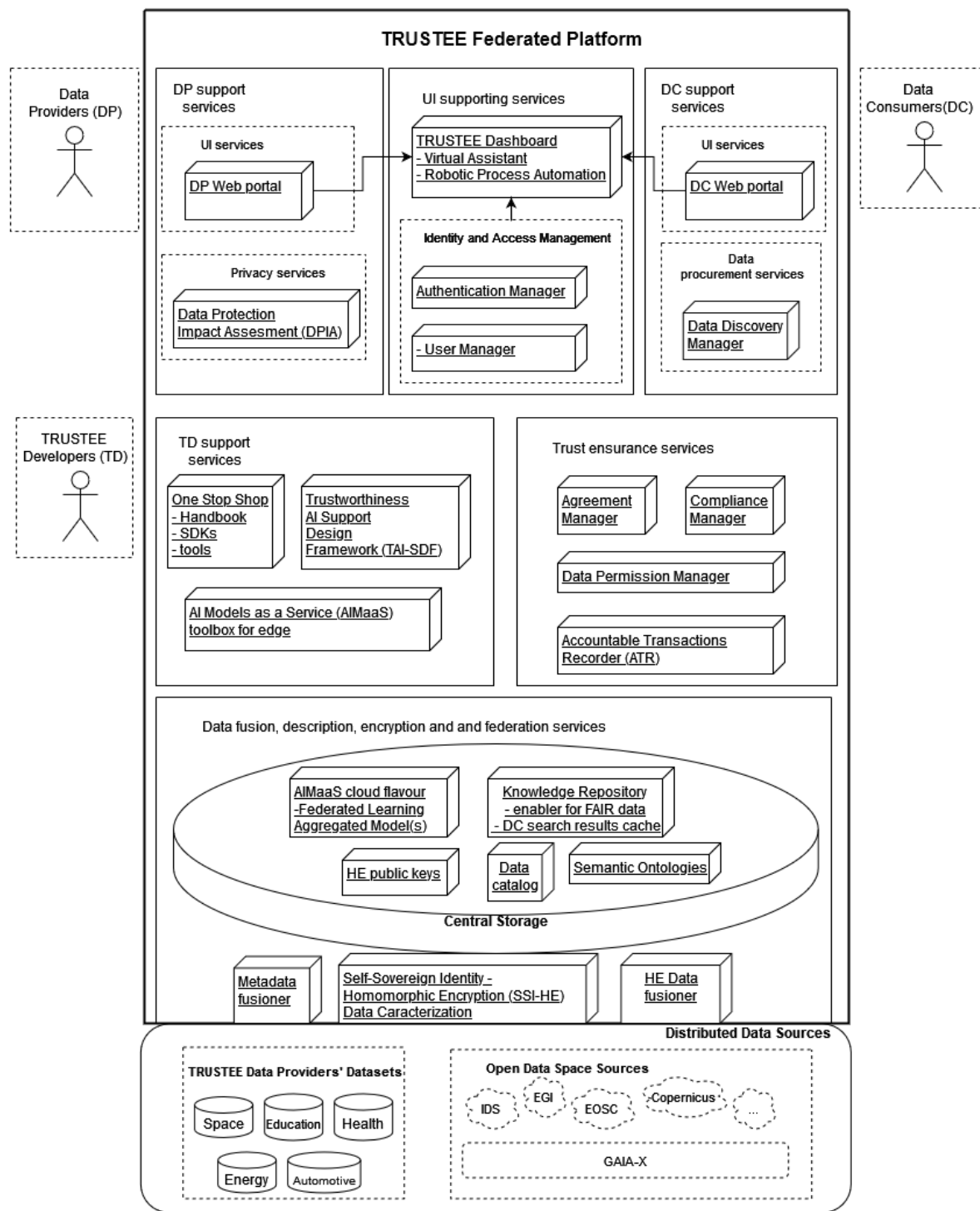


Figure 143: TRUSTEE conceptual architecture

ANNEX 1: INFORMATION CONCERNING THE PROCESSING OF THE DATA SHARED IN THE TRUSTEE SURVEY

The information contained in the following sections describes the details of the processing of your data obtained in the “Survey”, developed, and circulated for research purposes in our project, to understand end user requirements for the TRUSTEE platform. Please read it before providing the information to the survey and sharing the data with the TRUSTEE consortium.

1. The Project and Consortium

„Trust & Privacy Preserving Computing Platform For Cross-Border Federation Of Data“ Project - the TRUSTEE Project - is a project funded by the European Union’s Horizon EUROPE research & innovation programme under grant agreement No 101070214 and will last until 31 December 2025.

The TRUSTEE Project Activities have established aspiring goals to contend with a great challenge among EU countries: the improvement of the availability and quality of multidisciplinary big-data sources from various sectors in a trustworthy, fair, responsible and environmentally friendly manner, across the data lifecycle to enable data sharing and manipulation in compliance with prevailing and emerging legislation. It is TRUSTEE’s mission to combine technological and social innovations for secure and sustainable data operations (e.g., by optimizing/minimizing/de-centralizing processing, transfer and storage of data and avoiding unnecessary data manipulations, as well as technologies and solutions for fair and ethical collection, processing, and manipulation of these data), in line with the principles of responsible/trustworthy AI by using a co-development approach as backbone methodology.

The TRUSTEE consortium is listed below and incorporates the organizations from industry, academia, and research and development sector, who together conduct research:

Participant No.	Participant organisation name	Participant short name	Country
1 (CO)	IDRYMA TECHNOLOGIAS KAI EREVNAS	FORTH	Greece
2	ELLINIKO MESOGEIAKO PANEPISTIMIO	HMU	Greece
3	PANASONIC AUTOMOTIVE SYSTEMS EUROPE GMBH	PASEU	Germany
4	UNIVERSITÄT WIEN	UNIVIE	Austria

5	K3Y	K3Y	Bulgaria
6	ACCELIGENCE LTD	ACCELI	Cyprus
7	UNIVERSITA CATTOLICA DEL SACRO CUORE	UCSC	Italy
8	ERICSSON NIKOLA TESLA D.D.	ENT	Croatia
9	ADRESTIA EREVNITIKI IDIOTIKI KEFALAIIOUXIKI ETAIREIA	ADRESTIA	Greece
10	FUNDACION TECNALIA RESEARCH & INNOVATION	TECNALIA	Spain
11	RINA CONSULTING SPA	RINA-C	Italy
12	ZORTENET IDIOTIKI KEFALAIIOUXIKI ETAIREIA	ZORTENET	Greece
13	INQBIT INNOVATIONS SRL	IQB	Romania
14	TEKNOLOGIAN TUTKIMUSKESKUS VTT OY	VTT	Finland
15	FUJITSU TECHNOLOGY SOLUTIONS (LUXEMBOURG) SA	FUJITSU	Luxembourg
16	HEWLETT PACKARD ITALIANA SRL	HPE	Italy
17	HEWLETT-PACKARD CDS ITALIA SRL	HPE-CDS	Italy
18	INSTITUTUL DE STIINTE SPATIALE	ISS	Romania
19	CHECKWATT AB	CHECKWATT	Sweden

20	ATHINA-EREVNITIKO KENTRO KAINOTOMIAS STIS TECHNOLOGIES TIS PLIROFORIAS, TON EPIKOINONION KAI TIS GNOSIS	ATHINA	Greece
21	ETOILE PARTNERS LTD	EPL	UK
22	THE COURT OF EDINBURGH NAPIER UNIVERSITY	NAPIER	UK

For more information, please visit our website: <https://horizon-trustee.eu/>.

2. Purposes of data processing

The purpose of the TRUSTEE research project is the development and testing of a secure, trustworthy and privacy aware framework that will aggregate various interdisciplinary data repositories in the domains of Healthcare, Education, Energy, Space, Automotive and allow for cross-border communication between those data repositories.

As TRUSTEE is a research project, it will conduct a number of pilots in the mentioned domains, driven by the pilot use case leaders:

- CHECKWATT in the ENERGY domain;
- UCSC in the HEALTH domain;
- EPL in the EDUCATION domain;
- PASEU in the AUTOMOTIVE domain;
- ISS in the SPACE domain;
- VTT and ZORTENET in the cross-border/interdisciplinary data exchange.

To be able to understand the needs of the data exchange, data manipulation and data reuse in all those different domains that TRUSTEE could support, we need to collect information from the end users (among them, from you) about:

- The profile of TRUSTEE end user
- The information about the usual work of the end user (involving the data entry, structuring, manipulation, collection, modification, querying, sharing, analysis, and data computations), the amount of time spent on data use activities, current obstacles faced during these activities as well as the algorithms (AI and machine learning) that might be employed for aiding these processes.
- The expectations from the TRUSTEE solution and foreseen requirements, stakeholders, end-users, and type of use of this solution.

The data will be processed for TRUSTEE research purposes, especially in defining the structure of the TRUSTEE platform, the use cases scenarios, the legal and ethical framework for the development of the solution and for the data manipulation.

None of the personal data acquired will be disseminated or distributed outside the TRUSTEE consortium.

In the context of the TRUSTEE Project, your personal information will not be used for any other purposes or shared with any other party outside of those specified in this document.

3. Key Contacts

Project Coordinator: Emmanouil Spanakis <spanakis@ics.forth.gr>

Technical Manager: Evangelos Markakis <emarkakis@hmu.gr>

4. Data processing

a) What type of data will be processed?

The personal data categories which will be processed in the context of the study are as follows:

- Data related to your professional expertise (job role, years of experience)
- Age group
- Data related to typical duties in the workplace
- The name of the institution you are associated with
- Data use activities

Your data will be collected by the HMU and presented and transferred to the other TRUSTEE Consortium members in the form that will not allow your direct identification.

The data will be processed in accordance with Article 89 paragraph 1 of the European Union General Data Protection Regulation, by applying appropriate technical and organizational measures ensuring data minimisation.

b) Legal basis for the processing of your data

The legal basis under which your personal data shall be collected is in accordance with Article 6 paragraph 1 point (f) – legitimate interest of the TRUSTEE consortium, which is the performance of the research, collection and analysis of end-user requirements, the design and development of the TRUSTEE platform responding to the needs of end-user as well as testing and validation of the solution.

c) Who will be able to access your data?

No personal data will be shared with third parties, i.e. beyond the TRUSTEE consortium. The results of the research, which will be made available, will not contain any personal data.

TRUSTEE partners will have access to the data based on need-to-know principle during the lifecycle of the project. If a partner to the TRUSTEE consortium does not need to access the data for the purposes of the research, then no access should be granted to him/her.

d) Will your data be transferred to third countries outside of the European Union?

The data in the form that will not allow your direct identification may be transferred by HMU to the partners based in the United Kingdom (EPL and NAPIER). Nevertheless, please be ensured that those entities are obliged to adhere to the standards which are considered by the European Commission as adequate to the ones imposed by the EU data protection regulation, based on the adequacy decision adopted by the European Commission on 28 June 2021 (for further details see https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en).

e) For how long will the data be stored?

For the TRUSTEE project purposes, the data obtained from you will be stored for the lifetime of the project and five (5) years after the end of the project (according to TRUSTEE Grant Agreement, Data Sheet, Point 6). The project is foreseen to last until 31 December 2025. We will apply anonymisation and minimisation techniques in order to minimise any risk of confidentiality breach or unintentional data breach.

f) Is your data secure?

We do everything in our capacity to make sure that the confidentiality and security over the personal data we process is protected. We apply anonymization, pseudonymization and data minimization techniques. Access control and authentication-based environments are implemented to access the datasets containing personal data.

We also apply need-to-know principle when including any researcher being part of TRUSTEE team into personal data processing operation.

g) What are your rights?

You are hereby informed of your rights as a data subject in accordance with Article 12 to 14 of the GDPR.

You have the right to:

- Request information about whether we hold personal information about you, and, if so, what that information is and why we are holding it.
- Request access to your personal information. This enables you to receive a copy of the personal information we hold about you and to check that we are lawfully processing it.
- Request rectification of the personal information that we hold about you. This enables you to have any incomplete or inaccurate information we hold about you corrected.
- Request the restriction of processing of your personal information. This enables you to ask us to suspend the processing of personal information about you unless we are obligated to continue with the processing for compliance with an applicable statutory provision.
- Not to be subjected to automated decision-making including profiling. This right shall not apply if the automated decision making, and profiling is based on your explicit consent.
- You have the right to lodge a complaint against TRUSTEE consortium members regarding data protection issues with any data protection authority within the European Union (Article 13(2)(d) GDPR).
- Object. You have the right to object to the processing of your personal data, unless the processing is conducted on public interest grounds.

h) More information

You can find more information:

- In the Grant Agreement, which is made available to your institution as the party to the agreement.
- On our website and our social media.
- In the GDPR. Our main legal framework also includes the GDPR. Please have a look at it, especially at Article 89, which establishes minimum rules and procedures for data protection in the scientific research field.
- By asking us. We are more than happy to answer all questions you may have or, eventually, address you to our legal experts.

ANNEX 2: SURVEYS

2.1 PILOT LEADER

Partner acronym	

1. Role in the company

1.1. Please, define your job role in the company in which you are currently working (select the most relevant box)

Senior Researcher	☐
Associate Researcher	☐
Project Manager	☐
Technical Manager	☐
Software Developer	☐
Data Scientist	☐
Other	If your role does not fit to the choices above, please describe here:

2. Age group

Less than 30 years ☐	30 to 40 years ☐	40 to 50 years ☐	50 to 60 years ☐	More than 60 years ☐

3. Years of experience in your respective domain

Less than 3 years ☐	3 to 15 years ☐	More than 15 years ☐

4. A typical day in the job

4.1. Please, briefly describe your typical duties in your workplace

5. Data Usage Activities

5.1. Please, briefly describe your involvement in any data usage processes that you perform in your domain (data entry¹⁰⁰⁷, data manipulation¹⁰⁰⁸, data modification¹⁰⁰⁹, data querying¹⁰¹⁰, data sharing¹⁰¹¹, data analysis¹⁰¹², etc.)

5.2. Please, provide an indication of how much time do you spend in data usage activities.

Less than 1 hour ☐	2 to 4 hours ☐	4 to 6 hours ☐	6 to 8 hours ☐	More than 8 hours ☐

5.3. Please, indicate the obstacles you often come across regarding data usage activities

5.3.1. Data Entry

5.3.2. Data Manipulation

5.3.3. Data Modification

5.3.4. Data Querying

5.3.5. Data Sharing

5.3.6. Data Analysis

5.3.7. Other

5.4. Please, briefly describe the type of computations (data processing) you usually perform on data in your domain (mathematical operations¹⁰¹³, logical operations¹⁰¹⁴, visualization¹⁰¹⁵, etc.)

6. AI and Machine Learning

6.1. Please, imagine a realistic scenario where algorithms or automation (through AI and machine learning) may aid your work with data. For which data use and activities might algorithms or automation be most appropriate? Would you conduct any profiling¹⁰¹⁶ activities?

7. Data multi-disciplinarity and Secondary Use

7.1. Let's assume that data in your domain could be utilized for secondary use¹⁰¹⁷ purposes by other sectors using TRUSTEE. Moreover, let's assume that through TRUSTEE you could utilize data from other domains for secondary purposes. Other sectors could be education, healthcare, manufacturing, automotive, finance, space, energy, hospitality, climate, weather forecasting, transportation, etc. Please, reply to the best of your knowledge:

7.1.1. Which other sectors might benefit from **utilizing data from your domain**? (to whom might your data be useful)

¹⁰⁰⁷ Data entry refers to the insertion of data in a data platform

¹⁰⁰⁸ Data manipulation refers to performing logic or calculations to data in order to produce more refined data

¹⁰⁰⁹ Data modification refers to changing the actual values of the data through insertion, update, or deletion

¹⁰¹⁰ Data querying refers to requesting data or information stored in a repository

¹⁰¹¹ Data sharing refers to the exchange of data between multiple parties

¹⁰¹² Data analysis refers to the extraction of additional knowledge from the data itself

¹⁰¹³ Mathematical operations performed on data refer to calculations such as multiplication, addition, etc. on data values

¹⁰¹⁴ Logical operations performed on data refer to the use of logical functions, such as and, or, etc. (e.g., if A is true AND B is true, then C is false)

¹⁰¹⁵ Visualization of data refers to methods that assist in showcasing the differences and/or similarities between data

¹⁰¹⁶ Profiling means here any form of automated processing of personal data consisting of using data to evaluate certain personal aspects relating to a natural person, in particular, to analyze and predict aspects concerning that person's performance at work, economic situation, health, personal preferences, or interests, reliability, or behavior, location or movements. (Art. 4(4) GDPR)

¹⁰¹⁷ The term "secondary use of data" refers to the practice of putting collected information to work for a purpose that is distinct from their primary purpose of use

- 7.1.2. Which other data domains would be useful **to you**?
- 7.2. Please, describe at least two (or more) scenarios **of another sector** that might benefit from using the **data from your domain**
 - 7.2.1. Secondary Data domain that could utilize your data
 - 7.2.2. Their potential goal and benefits from utilizing your data
 - 7.2.3. Current obstacles from your side regarding providing access to secondary use of your data
- 7.3. Please, describe a clear example scenario of **your interest** in using **data coming from other domains**, for secondary purposes
 - 7.3.1. Secondary Data domain that you would be interested in
 - 7.3.2. Your goal and benefits that might be gained
 - 7.3.3. Current obstacles you foresee

8. Proportionality and necessity of processing

We understand that you will use fictitious or synthetic data in the TRUSTEE project. However, we would like to assess TRUSTEE platform from the perspective of the future end user, who may use real data on the platform. Assuming **the use of real data** in this scenario (which may be personal data), please reply to below questions.

- 8.1. Please indicate what **category**¹⁰¹⁸ of data would be potentially shared on the TRUSTEE platform by you in your domain and what kind of types of data you would like to access?
- 8.2. Please provide an indication of whether the processing purpose(s) are specified, explicit and legitimate.
- 8.3. What is your legal basis¹⁰¹⁹ making the processing lawful?
- 8.4. Is the processing of the data required to comply with a specific legal obligation under national law? If so, with which one?
- 8.5. Does the processing achieve your purpose?
- 8.6. Is there another way to achieve the same outcome? If personal data would be involved - could you achieve the same purpose with less of personal data?
- 8.7. How do you or would you ensure appropriate data quality in the context of described processing?
- 8.8. Are the data you process adequate, relevant, and limited to what is necessary for the purpose(s) you indicated? Are they accurate?
- 8.9. If you use AI/ machine learning for processing of personal data, how do you inform the data subjects about it and its use? How do you avoid bias?
- 8.10. What information will you give individuals?
- 8.11. Will you inform the data subjects about this type of processing? If yes, how?
- 8.12. How can data subjects exercise their rights (regarding the data sets provided onto TRUSTEE)?
- 8.13. Please indicate the potential benefits and risks that you envision that the described processing activity may create for or against end-users and the data subjects to whom processed data relate, if the data are personal.

9. TRUSTEE platform

- 9.1. Please, list your expectations from the TRUSTEE platform. What would you expect to be able to do through such a platform? What abilities would you like this platform to offer to you?
- 9.2. Please, indicate what you think are the most important requirements TRUSTEE should respect in your domain

¹⁰¹⁸ Any category of personal data (identifiers such a name, an identification number, location data, an online identifier, or one or more specific factors to the physical, psychological, genetic, mental economic, cultural or social identity of a person – race, ethnic origin, political opinions, religious or philosophical membership, genetic, health or biometric data, sex life, sexual orientation or others) and whether in pseudonymised form; or non-personal data, and if so, what category of non-personal data, not related to an identified or identifiable natural person (such as, e.g., the data about the legal entities, industrial data about product data of medical devices, vehicles, energy infrastructure).

¹⁰¹⁹ Please consider here one or more of the legal bases defined in Article 6(1) GDPR and, where the processing of special categories of data comes into play (which is prohibited if not based on a specific exception), please also indicate which exception defined in Article 9 GDPR applies.

9.3. Please, list several key stakeholders¹⁰²⁰ in your domain that would be interested in using TRUSTEE and their requirements for data privacy, security, interoperability, access to [what kinds of] data in common data spaces

9.4. Please, indicate possible end-users¹⁰²¹ that would be interested in using your data via TRUSTEE as well as their potential benefits and interests (please, identify **at least two** benefits and **two** interests)

10. Question from the Plenary

10.1. Please, indicate whether the main objectives of TRUSTEE are clear to you, or you require any additional explanation

11. Data analytics, legal and ethical considerations

11.1. Please, indicate who will participate in processing the data once it is made available on TRUSTEE

11.2. What provisions do you currently have for data security, and, where required, confidentiality (including technical and organizational measures, such as data recovery, secure storage/archiving, and transfer of sensitive data, etc.)?

11.3. Please, indicate whether informed consent for long-term preservation and secondary use of data may be needed once the data are made available on TRUSTEE. If you have examples, please attach them.

11.4. Do you currently deploy any algorithms to process your data?

11.4.1. Do you require consent from your data subjects for the deployment of algorithms for data processing? If now, why?

11.4.2. Please list the type of algorithms deployed. What role do they play in your environment?

11.5. Please, indicate whether the algorithms used at the moment (if any) produce legal effects or any similar effects significantly affecting the data subject/the user. If so, what kind?

11.6. What is considered high risk data manipulation in your domain and is any such done with the datasets that you will supply to TRUSTEE?

11.7. Do you interact with the natural persons to whom the data you would use on TRUSTEE Platform relate at some point and why? Do you collect the data directly from any data subject?

11.8. In your view, is there any reason to believe that at a certain point in the processing activities conducted by you the data should no longer be considered 'personal' (for example, by applying anonymization techniques)?

11.9. What mechanisms of oversight are implemented at pilots' ends to ensure that they respect individuals' digital rights?

2.2 PILOT MEMBER

Domain:

Please, select your sector from the table below:

Energy	☐
Healthcare	☐
Education	☐
Automotive	☐
Space	☐

¹⁰²⁰ Stakeholders might be companies, organizations etc. that could use the TRUSTEE platform in order to offer additional services to their clients, or stakeholders could be the individuals whose data could be processed by the TRUSTEE platform

¹⁰²¹ End-users might be individuals that could directly use the TRUSTEE platform

Other	If your domain / sector does not fit to the choices above, please describe here:

1. Role in the company

1.1. Please, define your job role in the company in which you are currently working (select the most relevant box)

Senior Researcher	☐
Associate Researcher	☐
Project Manager	☐
Technical Manager	☐
Software Developer	☐
Data Scientist	☐
Other	If your role does not fit to the choices above, please describe here:

2. Age group

Less than 30 years ☐	30 to 40 years ☐	40 to 50 years ☐	50 to 60 years ☐	More than 60 years ☐

3. Years of experience in your respective domain

Less than 3 years ☐	3 to 15 years ☐	More than 15 years ☐

4. A typical day in the job

4.1. Please, briefly describe your typical duties in your workplace

5. Data Usage Activities

5.1. Please, briefly describe your involvement in any data usage processes that you perform in your domain (data entry¹⁰²², data manipulation¹⁰²³, data modification¹⁰²⁴, data querying¹⁰²⁵, data sharing¹⁰²⁶, data analysis¹⁰²⁷, etc.)

5.2. Please, provide an indication of how much time do you spend in data usage activities.

Less than 1 hour ☐	2 to 4 hours ☐	4 to 6 hours ☐	6 to 8 hours ☐	More than 8 hours ☐

5.3. Please, indicate the obstacles you often come across regarding data usage activities

- 5.3.1. Data Entry
- 5.3.2. Data Manipulation
- 5.3.3. Data Modification
- 5.3.4. Data Querying
- 5.3.5. Data Sharing
- 5.3.6. Data Analysis
- 5.3.7. Other

5.4. Please, briefly describe the type of computations (data processing) you usually perform on data (i.e., data computations could be: mathematical operations¹⁰²⁸, logical operations¹⁰²⁹, visualization¹⁰³⁰, etc.)

6. AI and Machine Learning

6.1. Imagine a realistic scenario where algorithms or automation (through AI and ML) may help in your work. For which data use activities might algorithms or automation fit best?

7. Data multi-disciplinarity and Secondary Use

7.1. Let's assume that data in your domain could be utilized for secondary use¹⁰³¹ by other sectors through the "Platform". Simultaneously, through the "Platform" you could utilize data from other sectors for your own secondary purpose. Other sectors could be education, healthcare, manufacturing, automotive, finance, space, energy, hospitality, climate, weather forecasting, transportation, etc.

- 7.1.1. Which other sectors might benefit from **utilizing data from your domain**? (to whom might your data be useful) Which other data domains would be useful to you?
- 7.1.2. Which other sectors' data would be useful **to you**?

7.2. Please, describe a scenario **of another sector** that might benefit from using the **data from your sector**

- 7.2.1. Secondary Data sector that could utilize your data
- 7.2.2. Their potential goal and benefits from utilizing your data

¹⁰²² Data entry refers to the insertion of data in a data platform

¹⁰²³ Data manipulation refers to performing logic or calculations to data in order to produce more refined data

¹⁰²⁴ Data modification refers to changing the actual values of the data through insertion, update, or deletion

¹⁰²⁵ Data querying refers to requesting data or information stored in a repository

¹⁰²⁶ Data sharing refers to the exchange of data between multiple parties

¹⁰²⁷ Data analysis refers to the extraction of additional knowledge from the data itself

¹⁰²⁸ Mathematical operations performed on data refer to calculations such as multiplication, addition, etc. on data values

¹⁰²⁹ Logical operations performed on data refer to the use of logical functions, such as and, or, etc. (e.g., if A is true AND B is true, then C is false)

¹⁰³⁰ Visualization of data refers to methods that assist in showcasing the differences and/or similarities between data

¹⁰³¹ The term "secondary use of data" refers to the practice of putting collected information to work for a purpose that is distinct from their primary purpose of use

- 7.2.3. Current obstacles you experience in providing access to secondary use of your data, including necessary authorizations.
- 7.3. Please, describe a clear example scenario of **your interest** in using **data coming from other sectors**, for secondary purposes. Please tell us whether such data would potentially include any category of personal data¹⁰³².
 - 7.3.1. Secondary Data domain that you would be interested in
 - 7.3.2. Your goal and benefits that might be gained
 - 7.3.3. Current obstacles you foresee

8. Proportionality and necessity of processing

- 8.1. Please answer below questions with regards to the data processing activities which you indicated in points 5.1 and 6.1. Please try to provide the answers to the extent you can: Please provide an indication of whether the processing purpose(s) are specified, explicit and legitimate.
 - 8.1.1. Is the goal, for which you use the data, achieved through the processing that you described? e.g., to achieve the expected results of your research or to obtain any necessary information
 - 8.1.2. In case of a use of real, personal data, do you have your purpose always clearly specified and communicated to the people whose data you process?
 - 8.1.3. In the case of a processing described in point 6.1 – do you think there would be another way, not involving AI application, which would allow you to achieve the same outcome?
 - 8.1.4. How do you usually ensure appropriate data quality¹⁰³³ in the context of described data use from point 5.1 and 6.1?

9. Data Sharing Platform

- 9.1. Please, indicate whether you would be interested in using the “Platform” and **why** (you may find the description of the platform in the introduction of this document).
- 9.2. Please, list your expectations from the “Platform”. What would you expect to be able to do through such a platform? What abilities would you like this platform to offer to you?
- 9.3. Please, indicate what you think are the most important requirements the “Platform” should respect in your sector
- 9.4. Please indicate any category¹⁰³⁴ of personal data that would be potentially shared on the “Platform” by you in your sector

10. Data analytics, data protection and security

- 10.1. Do you deploy any algorithms to process your data currently?
 - 10.1.1. Please, list the type of algorithms deployed
- 10.2. What do you currently do to protect the data in your data use activities? Do you apply any kind of security measures?
- 10.3. Please answer the following questions only if you use personal data (or fictitious data resembling personal data):
 - 10.3.1. Do you collect the data directly from any person?

¹⁰³² Any category of personal data (identifiers such as a name, an identification number, location data, an online identifier, or one or more specific factors to the physical, psychological, genetic, mental economic, cultural or social identity of a person – race, ethnic origin, political opinions, religious or philosophical membership, genetic, health or biometric data, sex life, sexual orientation or others) and whether in pseudonymised form; or non-personal data, and if so, what category of non-personal data, not related to an identified or identifiable natural person (such as, e.g., the data about the legal entities, industrial data about product data of medical devices, vehicles, energy infrastructure).

¹⁰³³ e.g., how do you ensure that the data are adequate, relevant, and limited to what is necessary

¹⁰³⁴ Any category of personal data (identifiers such as a name, an identification number, location data, an online identifier, or one or more specific factors to the physical, psychological, genetic, mental economic, cultural or social identity of a person – race, ethnic origin, political opinions, religious or philosophical membership, genetic, health or biometric data, sex life, sexual orientation or others) and whether in pseudonymised form; or non-personal data, and if so, what category of non-personal data, not related to an identified or identifiable natural person (such as, e.g., the data about the legal entities, industrial data about product data of medical devices, vehicles, energy infrastructure).

- 10.3.2. Would you conduct any profiling¹⁰³⁵ activities?
- 10.3.3. Do you consider the data which you use in your data use activities (as described in section 5) as anonymous data¹⁰³⁶? If so, and if possible for you, please explain why you think so.
- 10.3.4. Do you know what benefits or risks the daily data use you described in point 5.1, could bring to the people whose data you process?
- 10.3.5. Do you know what benefits or risks of a potential use of AI models, as described in point 6.1, could bring to the people whose data you process?

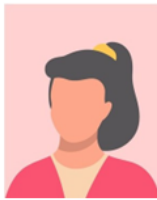
¹⁰³⁵ Profiling means here any form of automated processing of personal data consisting of using data to evaluate certain personal aspects relating to a natural person, in particular, to analyze and predict aspects concerning that person's performance at work, economic situation, health, personal preferences, or interests, reliability, or behavior, location or movements. (Art. 4(4) GDPR)

¹⁰³⁶ Anonymised data - information which does not relate to an identified or identifiable natural person or personal data rendered anonymous in such a manner that the person is not or no longer identifiable. If re-identification is not reasonably likely that someone would be able to do it (e.g., if it would be very time- and cost-consuming and would require application of very advanced technology that is non-existing or barely used nowadays), then that the data can be considered anonymised.

3.1 ENERGY DOMAIN

PERSONAS

DATA PROVIDER

**Glenda Williams**

Domain: Energy
Age group: 60+ years
Occupation: Project Manager
Years of experience: 15+ years
Daily data usage: 2-4 hours
TRUSTEE Role: Data Provider

"I want secure sharing of data for secondary use"

Goals

- Secure data manipulation and sharing on the cloud
- Give secondary use data for optimization research, energy prediction and consumption models
- Give data for secondary use to academia, electricity/power intensive industry, smart home, energy, weather forecasting, climate, transportation, manufacturing, education, and finance domains
- Secure and efficient data sharing with partners
- Readily available and ease-of-use data

Frustrations

- "It frustrates me that there are non-standardized methods or defined validation functions for data entry, and unclear instructions for correct data formatting"
- "It frustrates me that some data cannot be shared due to privacy and/or business policies"
- "It frustrates me that cyber security attacks result in taking control of big energy resources"
- "It frustrates me that the sharing of the data becomes problematic when the data needs to be transformed/reformatted"
- "It frustrates me that it is difficult to provide data for secondary use due to issues relating to data collection, sharing business model, marketing the availability of data, and automatic setup of data flows"

Glenda is a project manager in the energy domain. She ensures data quality with manual evaluation, automatic evaluation as part of automatic data collection routines, technical measures, as well as by directly retrieving them from reliable and primary sources. She validates appropriate data by taking a data sample and applying calculations, as well as the use of visualisation platforms to test and validate the data. She can share consumption and production data together with user metadata, location data, and data about energy production and availability of wind power turbines, owned by small businesses or individuals. She believes that electricity/power intensive industry might use data from her sector to curtail their costs and make them more profitable. Researchers who want to develop new algorithms for energy prediction, optimization, and consumption could utilize her data to improve local weather forecasts and analyse behavioural patterns. The energy market can use availability of wind power and production data for forecasting, so that they can produce more stable and predictable market prices. Education could utilize her data to study energy production/consumption both for household and industrial buildings. Manufacturing could utilize her data to learning the ways hardware is used in the energy systems to give the manufactory an indication and better understanding on improvement areas. Finance could utilize her data to offer users better deals based on their energy consumption and habits. Climate could utilize her data to create awareness and encourage the user to reduce their climate footprint. The energy related habits and goals could help a smart home system make it smarter and more energy efficient. She is interested in a platform where she will be able to share commercial and private data with other entities without revealing details that are business secrets or part of privacy data. Furthermore, she wants this platform to enable her new cooperations at a market level. She believes that it should allow her to be able to control and limit access to secret/private data, as well as be able to control who have access rights to shared data and identify and use shared data from external sources.

CONSUMER



Harry Walton

Domain: Energy

Age group: 60+ years

Occupation: Senior Researcher

Years of experience: 15+ years

Daily data usage: 6-8 hours

TRUSTEE Role: Consumer

Goals

- Get data for secondary use from education, weather forecasting, climate, environmental, finance, space, energy, international commerce and natural resources related domains
- Secure and efficient energy data sharing
- Get access to relevant data sources that can be used as internal data
- Understand how an AI/ML model interprets the data I am interested in
- Evaluate the compliance of AI models with current standards, policies and regulations
- Finetune the process of Data Analysis while using different algorithms without the need to be an expert
- All available data in one open and free ecosystem

"I want easier access to available data"

Frustrations

- "It frustrates me that mandatory changes of file formats are time consuming"
- "It frustrates me that there is a lack of data availability due to security and privacy concerns in data manipulation and querying"
- "It frustrates me that querying and retrieving data from critical and live databases requires additional expertise"
- "It frustrates me that many data are not available or free to access"
- "It frustrates me that I cannot access data for secondary use due to quality of data, data not shared in a granular manner, and only highly aggregated data"
- "It frustrates me that there is lack of interoperability and integration between different data sources due to privacy concerns in data modification and the lack of standards compliance in data sharing"
- "It frustrates me that data analysis may produce errors due to low data quality and missing data/meta-data"
- "It frustrates me that it is difficult to validate the results manipulation of data algorithms"
- "It frustrates me that in data analysis it is challenging to test and present results in a user-friendly way"
- "It frustrates me that different systems require different algorithms to manipulate data"
- "It frustrates me that I'm unable to process and manipulate the data format"

Harry is a Senior Researcher in the energy domain. His typical day at work includes project definition, technical specifications, requirements engineering, and company management duties. He is interested to access education data to learn and improve his calculations/algorithms based on new studies. Finance data can keep track of the energy service business area. Climate data can keep track of studies and news in order to be informed, aware, and share it with customers. Weather forecasting data can enhance his work with solar and wind energy. Space data, as well as climate and environmental data could also be useful to him and assist researchers to better understand solar panel performances and forecasting productions. Data from international commerce and data related to natural resources demand, supply and availability could be useful to him. Also, data about long and medium term energy price forecasts can be used for his optimal production. Natural resource data could be useful to determine the market availability and demand of resources to forecast the market trends. He is interested in AI/ML algorithms for selecting and deploying the best-suited visualization depending on the result data, and accurately forecasting various types of time-series data and trends in time-series data, for data analysis, finding patterns, making predictions, comparing equipment, for automatic validation of querying input data, automatic validation of data modification and manipulation tests, automatic quality validation of data, automatic running of tests, and benchmark installations. He is interested in a platform where he will be able to share commercial and private data with other entities without revealing details that are business secrets or part of privacy data. Furthermore, he wants this platform to enable him new cooperations at a market level. He also believes that the platform should allow him to be able to control and limit access to secret/private data, as well as be able to control who have access rights to shared data and identify and use shared data from external sources.

MODEL PROVIDER



Marsha Mills

Domain: Energy

Age group: <30 years

Occupation: Data Analyst

Years of experience: 3 - 15 years

Daily data usage: 2 - 4 hours

TRUSTEE Role: Model Provider

Goals

- Secure and fast data manipulation and sharing on the cloud
- Standardized interoperability compliant with all security requirements and legislations

Frustrations

- "It frustrates me that I need to use complicated systems for simple data entry"
- "It frustrates me that the high volumes of Real World Data require improved analysis methods and tools"
- "It frustrates me that there are errors in data manipulation due to the low quality and inconsistency of input data formats and the non-standardized decimal conversions"
- "It frustrates me that data analysis may produce errors due to low data quality or missing data/meta-data"
- "It frustrates me that querying and retrieving data from critical and live databases requires additional expertise because different sources may have different query interfaces and tools"
- "It frustrates me that mandatory changes of file formats are time consuming"

"I want to access available data and perform analysis on them"

Marsha is a Data Analyst in the energy domain. Her typical day at work includes preliminary analysis using publicly available data and business case and technical simulations using private customer data. The data processes that she is involved in include modification, querying, analysing, and sharing energy data from customers, retrieving and analysing financial data at company and project level, and entering budget data. The types of computations she usually performs include mathematical operations for grid data, logical operations for the system, visualization of the data in different forms, and format transformations for detection of cyber threats. Furthermore, she utilizes algorithms to group and categorize customers based on metadata and algorithms for analysis of logs for anomaly detection related to cyber threats and equipment status analysis. She would be interested in a platform where she could be able to process data fast and share energy data using a secure cloud environment. She wants the rich-data platform to be easy to use, open, free, and contain good inbuilt visualization tools. A platform where she can have access to different types of data to test and find new interesting information that can provide to the users or use to improve their systems.

USER STORIES

ID	User Story

DATA PROVIDER	
DP22	As a Project Manager in the energy domain, I want to be confident that cyber security attacks will not result in taking control of big energy resources, so that I can easily and securely share my data for optimization research, energy prediction, and consumption models
DP23	As a Project Manager in the energy domain, I want to use an efficient and highly secure platform, so that I can share my energy data with others
DP24	As a Project Manager in the energy domain, I want to use standardized methods and clear instructions for correct data entry and formatting, so that I can make my data readily available and share them without reformatting
DP25	As a Project Manager in the energy domain, I want to easily transform/reformat energy data, so that I can share them with the academia and electricity/power intensive industry for secondary use
DP26	As a Project Manager in the energy domain, I want to overcome privacy and business policies, so that I can share data with partners
DP27	As a Project Manager in the energy domain, I want to overcome the issues relating to data collection, sharing business model, marketing the availability of data, and automatic setup of data flows, so that I can give data for secondary use to energy, weather forecasting, transportation, manufacturing, and finance domains
DP28	As a Project Manager in the energy domain, I want to overcome privacy regulations, so that I can give data for secondary use to education, manufacturing, finance, climate, and smart home domains
DP29	As a Project Manager in the energy domain, I want define functions to validate the entry data easier, so that I can enter, validate, and share data faster
CONSUMER	
C22	As a Senior Researcher in the energy domain, I want to access a platform that is interoperable and integrates data of different data sources without the need of aligning different data formats, so that I can access available and free data in one ecosystem and put more effort in data computations
C23	As a Senior Researcher in the energy domain, I want to access climate, environmental, and space data for secondary use, so that I can advance my research regarding optimization, energy prediction, and consumption models
C24	As a Senior Researcher in the energy domain, I want to have access in higher volumes of data and extract knowledge without the need for me to be an expert in complex analysis methods and tools, so that I can advance my research outcomes in a more fast and efficient way

C25	As a Senior Researcher in the energy domain, I want to have access to a standardized interoperability platform, so that I can access and modify data easily and fast without the time overhead of changing of file formats
C26	As a Senior Researcher in the energy domain, I want to have access to a platform compliant with all security requirements and legislations regarding data manipulation and querying, so that I can access and use a variety of available data
C27	As a Senior Researcher in the energy domain, I want to use data that contain meta-data, so that my analysis and simulations become more robust and less erroneous
C28	As a Senior Researcher in the energy domain, I want to access data from critical and live databases without having to become a database expert, so that I can expand my research faster
C29	As a Senior Researcher in the energy domain, I want to overcome the issues relating to quality of data, data not shared in a granular manner, and availability of only highly aggregated data, so that I can access data for secondary use from weather forecasting, finance, and energy domains
C30	As a Senior Researcher in the energy domain, I want to use a platform where commercial and private data can be shared securely, so that I can access relevant data sources that can be used as internal data
C31	As a Senior Researcher in the energy domain, I want to access data about long- and medium-term energy price forecasts, so that I can use them for optimal production
C32	As a Senior Researcher in the energy domain, I want to access education data, so that I can learn and improve my calculations/algorithms based on new studies
C33	As a Senior Researcher in the energy domain, I want to access finance data, so that I can keep track of the energy service business area
C34	As a Senior Researcher in the energy domain, I want to access climate data, so that I can keep track of studies and news in order to be informed, aware, and share it with customers
C35	As a Senior Researcher in the energy domain, I want to access weather forecasting data, so that I can further expand my solar and wind energy research
C36	As a Senior Researcher in the energy domain, I want to have access to improved AI-based analysis methods and tools which I can easily interpret without investing a huge amount of time to understand every process, so that I can improve my data analysis, finetune the process, and extract knowledge faster and efficiently

C37	As a Senior Researcher in the energy domain, I want to be able to assess the compliance of an AI model to policies, regulations, and standards, so that I can trust its results and use them without any legal or ethical issues
C38	As a Senior Researcher in the energy domain, I want to Utilize AI/ML algorithms for selecting and deploying the best-suited visualization depending on the result data, and accurately forecasting various types of time-series data and trends in time-series data, for data analysis, finding patterns, making predictions, comparing equipment, for automatic validation of querying input data, automatic validation of data modification and manipulation tests, automatic quality validation of data, automatic running of tests, and benchmark installations, so that I can further expand my research
C39	As a Senior Researcher in the energy domain, I want to access natural resource data and perform data analysis to determine market availability and demand of resources, so that I can forecast market trends
C40	As a Senior Researcher in the energy domain, I want to be able to validate any data manipulation results, so that I can validate and enhance my research outcomes faster
C41	As a Senior Researcher in the energy domain, I want to test and present the data analysis results in a user-friendly way, so that I can better illustrate and showcase my research outcomes
C42	As a Senior Researcher in the energy domain, I want to utilize the same algorithms to manipulate different types of data and use a single platform for this manipulation, so that I can produce scientific results faster
C43	As a Senior Researcher in the energy domain, I want to be able to process and manipulate the format of data, so that I can produce scientific results faster
MODEL PROVIDER	
MP22	As a Data Analyst in the energy domain, I want to eliminate the mandatory change of file formats, so that I can manipulate my data efficiently and fast
MP23	As a Data Analyst in the energy domain, I want to use simple systems for data entry, so that I can enter data easier and faster
MP24	As a Data Analyst in the energy domain, I want to manipulate data with consistent input formats and decimal conversions, so that I can spend more effort and time in data processing
MP25	As a Data Analyst in the energy domain, I want to manipulate and analyze good quality data without missing values, so that I can advance my research in ML and AI algorithms development for the energy domain
MP26	As a Data Analyst in the energy domain, I want to manipulate good quality data without missing values, so that I can be able to perform data analysis on them using my developed AI and ML models

MP27	As a Data Analyst in the energy domain, I want to use similar interfaces and tools to query data from multiple sources, so that I do not need to get familiar with different approaches every time I want to access data from different domains
MP28	As a Data Analyst in the energy domain, I want to provide algorithms to group and categorize customers based on metadata, and algorithms for analysis of logs for anomaly detection related to cyber threats and equipment status analysis, so that I can assist researchers in other domains

3.2 HEALTHCARE DOMAIN

PERSONAS

DATA PROVIDER



Christy Harvey

Domain: Healthcare

Age group: 60+ years

Occupation: Technical Manager

Years of experience: 15+ years

Daily data usage: 2-4 hours

TRUSTEE Role: Data Provider

Goals

- Give data for secondary use to insurance companies, simulation, planning, security about interaction with machines, food market, health planning domains, finance, healthcare, hospitality, nutrition, and physical activity domains
- Exchange data between manufacturing, finance, education, administrative, and security domains
- Operate on data without having to pseudonymize them

Frustrations

- "It frustrates me that there are no secure platforms with waivers from local IRB"
- "It frustrates me that sometimes data are not easy to be inserted and there is no option to add extras"
- "It frustrates me that I have to manually anonymize data"
- "It frustrates me that sharing data is difficult due to cybersecurity issues and personal data protection regulations"
- "It frustrates me that sometimes there is a lack of unequivocal labels"
- "It frustrates me that it is time consuming to evaluate data entries of users"
- "It frustrates me that there is lack of referential integrity in data modification"
- "It frustrates me that I have to perform manual change of data formats"

"I want to exchange data for research"

Christy is a Technical Manager in the healthcare domain. She supports that education manufacturing domains can utilize her data to produce early simplified diagnosis with limited resources, and to understand the relationship between well-being and knowledge. She can provide non-personal pharmacological data and data from medical devices, as well as personal identifier, physical, psychological, genetic, health or biometric pseudonymized data. She can share clinical, demographic, surgical, medical, comorbidity, pathologic, histological, and treatments pseudonymized data. She believes that a health company could be interested in analysing x-ray machine data to improve or create dedicated products in the field of radiology, in order to have real parameters to use in their simulations to improve device performances. Furthermore, nutrition domains could utilize her data to improve Nutritional Screening or creating a standardized Nutritional Program to help those in need of support. She also believes that insurance companies and the food market could benefit from her data because they can target the average of adverse events like diseases and predict clinical scenarios or study the risk of medical treatments and calculate their costs. She would like to exchange data with the education domain for healthcare planning, provisioning of services, and development of regional and national policies. She is interested in a rich-data platform where she can input and access data from other hospitals divided by areas of interest with a precise description of the characteristics of the data so that she can be sure of their consistency. She is interested to use a platform where she could work on data without having to manually pseudonymize them. She would also like to utilize a platform that is free, easy-to-use, secure, offers statistical tools, accept data in different formats with unequivocal labels, supports transcription patient identification data with second automatic match to avoid transcription errors, and supports anonymization of data.

CONSUMER



Gregory Nash

Domain: Healthcare
Age group: 50 – 60 years
Occupation: Senior Researcher
Years of experience: 3 – 15 years
Daily data usage: 2 – 4 hours
TRUSTEE Role: Consumer

Goals

- Get data for secondary use from healthcare-related companies, facilities, nutrition, diet, sport activities, multi-omics, clinical platforms, domestic devices, and quality driving car systems related domains
- Exchange data between manufacturing, finance, education, administrative, security, and local community social/building/service management domains
- Access files registered in different health systems
- Use of trustworthy AI/ML models that support all elements of trustworthiness such as compliance, fairness, explainability, robustness, security, and privacy

“I want to utilize a rich-data platform”

Frustrations

- “It frustrates me that there are a lot of categorical data but little data with continuous distribution”
- “It frustrates me that I cannot be informed about the available data”
- “It frustrates me that I often have to identify the optimal queries which is not always simple, nor immediate”
- “It frustrates me that knowledge of programming languages is required to perform complex data queries”
- “It frustrates me that there are no easy-to-use filters in data querying”
- “It frustrates me that data querying is limited to some functions”
- “It frustrates me that there are no methods to check consistency in data entry”
- “It frustrates me that often the data provided are not of high quality and sometimes they are not even valid”
- “It frustrates me that it is time consuming to evaluate data entries of users”

Gregory is a Senior Researcher in the healthcare domain. He finds data stemming out from domains relating to pharmaceutical companies, local government organizations, and healthcare facilities useful for research. He could access data from domestic devices and quality driving car systems to better monitor the patient's status and improve security in case of diseases. He would like to utilize data from finance and education to model data and observe them, and to educate citizens on healthy lifestyle, physical activities and nutrition. He wants to use secondary data from multi-omics platforms or clinical platforms to create surrogate arms in trials so that he can predict responses to oncological therapies and overcome difficulties in running and developing surgical trials, as well as their cost effectiveness. He would like to exchange data with the education domain for healthcare planning, provisioning of services, and development of regional and national policies. He is curious about Machine Learning algorithms for patterns identification in Case Record forms. He is interested in AI/ML algorithms for implementation of clinical pathways and healthcare management, prediction of adverse events or risk of treatments and surgery times or risks, benchmarking activities, analysis of evolutive dynamics in the clinical course based on available information, decision support systems in surgeries, image analysis, automated diagnosis, radiomic and genomic data, prediction of accuracy of big datasets, prevention of AE and SAE, and pseudonymization techniques. Furthermore, he could utilize data from manufacturing devices domains for monitoring vital parameters to implement, improve, and simplify clinical monitoring. He is also interested in a rich-data platform where he can input and access data from other hospitals divided by areas of interest with a precise description of the characteristics of the data so that he can be sure of their consistency. He would also like for this platform to be easy to use, and he would be interested in accessing numerical and descriptive data of clinical parameters.

MODEL PROVIDER



Tomas Fleming

Domain: Healthcare
Age group: 30 – 40 years
Occupation: Data Scientist
Years of experience: 3 – 15 years
Daily data usage: 6 – 8 hours
TRUSTEE Role: Model Provider

Goals

- Compare data from different hospitals that use the same technologies

“I want to modify and process data fast and easy”

Frustrations

- “It frustrates me that there are not any implemented user-friendly statistical tools”
- “It frustrates me that I cannot automatically check consistency of data from different sources”
- “It frustrates me that the availability and sustainability of health-related data is costly”
- “It frustrates me that data entry is manual and therefore is difficult to avoid errors”
- “It frustrates me that I cannot automatically check consistency of data from different sources”
- “It frustrates me that data entry is difficult and time consuming and most students or data managers find it hard and eventually do not perform it”
- “It frustrates me that the collection of health-related data is difficult”

Tomas is a Data Scientist in the healthcare domain. His typical day in the work includes uploading of personal and health data to different software for research and analysis of these data. His main data use processes relate to data modification and querying by utilizing, mainly, logical operations. The computations he typically performs include statistical analysis, mathematical and logical operations, and visualization on data, such as tables, texts, or heatmaps. He would like to be able to compare data and perform computations on them without having performance overhead or increasing data storage. He is interested in a platform that is free, easy-to-use, secure, offers statistical tools, accept data in different formats with unequivocal labels, supports transcription patient identification data with second automatic match to avoid transcription errors, and supports anonymization of data.

USER STORIES

ID	User Story

DATA PROVIDER	
DP44	As a Technical Manager in the healthcare domain, I want to encrypt data easily without having performance overhead or increasing data storage, so that I can avoid pseudonymization and share data faster
DP45	As a Technical Manager in the healthcare domain, I want to use a secure platform with waivers from local IRB and privacy policies, so that I can share data between finance, education, administrative, and security related domains
DP46	As a Technical Manager in the healthcare domain, I want to give data for secondary use to insurance companies, simulation, planning, security about interaction with machines, and food market domains, so that I can share knowledge extracted from healthcare data
DP47	As a Technical Manager in the healthcare domain, I want to use a platform that accepts data in different formats with unequivocal labels, and can automatically anonymize data, so that I can share data with other domains easily
DP48	As a Technical Manager in the healthcare domain, I want to be able to automatically add data and include extra metadata fields to these data if needed, so that I can perform data entry faster and in a more complete manner while avoiding human errors
DP49	As a Technical Manager in the healthcare domain, I want to give data for secondary use to education, finance, nutrition, physical activity, healthcare, hospitality, and health planning related domains, so that I can assist in the advancement of scientific research
DP50	As a Technical Manager in the healthcare domain, I want to use a secure platform free of cybersecurity and personal data privacy issues, so that I can share and exchange data with other domains
DP51	As a Technical Manager in the healthcare domain, I want to automatically evaluate data entries of users, so that I can import data faster
DP52	As a Technical Manager in the healthcare domain, I want to have more referential integrity in data modification and perform automatic changes in data formats, so that I can operate on valid data faster without the need to manually modify them
CONSUMER	
C44	As a Senior Researcher in the healthcare domain, I want to be able to perform complex queries without the knowledge of a programming language, so that I can find the data I want easier, faster, and more efficiently
C45	As a Senior Researcher in the healthcare domain, I want to access data of the education domain, so that I can research on healthcare planning, provisioning of services, and development of regional and national policies

C46	As a Senior Researcher in the healthcare domain, I want to access secondary data from multi-omics or clinical platforms, so that I can predict responses to oncological therapies and overcome difficulties in running and developing surgical trials, as well as their cost effectiveness
C47	As a Senior Researcher in the healthcare domain, I want to access data from domotic devices, quality driving car systems, nutrition, diet, and sport activities, so that I can monitor the patient's status and improve security in case of diseases
C48	As a Senior Researcher in the healthcare domain, I want to utilize a rich-data platform where I can access data from other hospitals divided by areas of interest and with a precise description of their characteristics, so that I can be sure of the data consistency
C49	As a Senior Researcher in the healthcare domain, I want to use a platform that is user-friendly and enables me to access numerical and descriptive data, so that I can further expand my research and knowledge expertise
C50	As a Senior Researcher in the healthcare domain, I want to utilize easy-to-use filters in data querying, so that I can perform analysis on those data faster without expert's knowledge on programming languages
C51	As a Senior Researcher in the healthcare domain, I want to access data from various domains, such as finance and education, so that I can model data and extract knowledge for educating citizens on healthy lifestyle, physical activities, and nutrition
C52	As a Senior Researcher in the healthcare domain, I want to utilize a platform that offers multiple querying functions, as well as automatic identification of the optimal queries, so that I can query data easier and faster
C53	As a Senior Researcher in the healthcare domain, I want to access files and data registered in different hospitals and healthcare systems, so that I can compare the use of same and different technologies
C54	As a Senior Researcher in the healthcare domain, I want to access data from local community social/building/service management, so that I can expand my research and extract further knowledge about disease outburst in certain areas, communities, etc.
C55	As a Senior Researcher in the healthcare domain, I want to access data with continuous distribution, so that I can expand my research using a variety of data compared to the categorical data that are mostly shared in data repositories
C56	As a Senior Researcher in the healthcare domain, I want to be able be able assess the validity, quality, and fairness of the data I access, so that I can choose which of them I will utilize in my research in order to produce optimum results

C57	As a Senior Researcher in the healthcare domain, I want to be able to be informed about the available data that I can access, so that I can utilize them in my research and produce better results
C58	As a Senior Researcher in the healthcare domain, I want to use a platform that offers the ability to use fully trustworthy AI/ML models, so that I can produce high quality results that conform with current regulations, standards, and policies
C59	As a Senior Researcher in the healthcare domain, I want to utilize AI/ML algorithms for patterns identification in Case Record forms, implementation of clinical pathways and healthcare management, prediction of adverse events or risk of treatments and surgery times or risks, benchmarking activities, analysis of evolutive dynamics in the clinical course based on available information, decision support systems in surgeries, image analysis, automated diagnosis, radiomic and genomic data, prediction of accuracy of big datasets, prevention of AE and SAE, and pseudonymization techniques, so that I can further expand my research
C60	As a Senior Researcher in the healthcare domain, I want to automatically evaluate data entries of users, so that I can perform statistical analysis faster without errors
C61	As a Senior Researcher in the healthcare domain, I want to utilize data from healthcare devices for monitoring vital patient parameters, so that I can implement, improve, and simplify clinical monitoring

MODEL PROVIDER

MP44	As a Data Scientist in the healthcare domain, I want to be able to sustain collected health-related data easier and at lower costs, so that I can use and re-use healthcare data
MP45	As a Data Scientist in the healthcare domain, I want to use methods to check consistency in data entry, and make data entry easier and faster, so that I can assign it to students or data managers
MP46	As a Data Scientist in the healthcare domain, I want to utilize a platform that is free, easy-to-use, secure, offers user-friendly statistical tools, so that I can manipulate and analyse my data easier and faster
MP47	As a Data Scientist in the healthcare domain, I want to automatically check consistency of data from different sources and automatically fix the mistakes in these data, so that I can perform data manipulation faster and save time

3.3 EDUCATION DOMAIN

PERSONAS

DATA PROVIDER



Saul Lucas

Domain: Education
Age group: 40-50 years
Occupation: Senior Researcher
Years of experience: 15+ years
Daily data usage: <1 hour
TRUSTEE Role: Data Provider

"I want to securely share data with others"

Goals

- Give secondary data to energy and other domains that evaluate budget and consumption rates
- Provide the means to data subjects to exercise their rights, requests, and interests regarding their data
- Secure transfer and sharing of data with other institutions

Frustrations

- "It frustrates me that there are high risks in data sharing regarding data loss and privacy"
- "It frustrates me that there are many legal and ethical limitations in education data sharing"
- "It frustrates me that there is high risk of education data misuse that aims at profiling students"
- "It frustrates me that students are automatically subjected to laws and conditions of companies that their institution forces them to use"

Saul is a Senior Researcher in the education domain and his daily routine includes collecting data from students for research purposes, such as course attendance and performance, as well as entering data into virtual learning environments, and sharing data with other faculty members. He is involved in data use processes relating to end-user privacy, digital and human rights, and various EdTech applications in education settings. Although he supports data sharing, he is sceptical about the data loss and privacy risks that data sharing may include. He is a strong supporter of protecting the students' data privacy and he believes that education data could be utilized by various organizations and governmental institutions looking at budgets, expenditure in EdTech, learning interventions, and energy consumption rates. He believes that the data he is collecting need to be used not only for educational purposes but also for research. He can share personal, psychological, political opinions, and purchasing data. He would be interested in a platform that will enable data transfer and sharing with other institutions while supporting privacy, protection from surveillance, cross platform functioning, as well as ease of use.

CONSUMER



Eleanor Dennis

Domain: Education
Age group: 30-40 years
Occupation: Researcher
Years of experience: 3-15 years
Daily data usage: 6-8 hours
TRUSTEE Role: Consumer

"I want to be able to use authenticated and available data"

Goals

- Use secondary data from economy, job market, energy, and health domain
- Secure transfer and sharing of data with other institutions
- Data sharing for secondary use with education and healthcare domain
- Support of data privacy and protection from surveillance
- Evaluation through AI of the way a website/portal collects and/or uses the collected data
- AI analysis of data that will provide information on specific interventions and learning outcomes in the near future

Frustrations

- "It frustrates me that there are high risks in data sharing regarding data loss and privacy"
- "It frustrates me that there are many legal and ethical limitations in education data sharing"
- "It frustrates me that it is difficult to confirm the authenticity of data"
- "It frustrates me that many organizations don't share their data for secondary use"
- "It frustrates me that it is difficult to access and use data that are provided across different platforms"
- "It frustrates me that validating AI support decisions is usually impossible"

Eleanor is a Researcher in the education domain. Her typical duties may include research, participation in conferences and meetings, consultancy, and teaching. Although she supports data sharing, she is sceptical about the data loss and privacy risks that data sharing may include. She is a strong supporter of protecting the students' data privacy and she believes that education data could be utilized by various organizations and governmental institutions looking at budgets, expenditure in EdTech, learning interventions, and energy consumption rates. She thinks that an AI model that could collect information from different websites/portals and track the use of her data would be very interesting and useful. Furthermore, she foresees that the education domain would benefit from data provided by economy, job market, energy, and healthcare domains. She would be interested in a platform that will enable data transfer and sharing with other institutions while supporting privacy, protection from surveillance, cross platform functioning, as well as ease of use.

MODEL PROVIDER



Emelia Hood

Domain: Education
Age group: 50-60 years
Occupation: Data Scientist
Years of experience: 3-15 years
Daily data usage: 4-6 hours
TRUSTEE Role: Model Provider

"I want AI support and cross platform functioning"

Goals

- Focus on operational data for sustainability
- Cross platform functioning
- Creation of personal account

Frustrations

- "It frustrates me when technical solutions are not easy to use and require expert's knowledge"

Emelia is a Data Scientist in the education domain. The types of computations she usually performs on data relates to evaluating risks and mapping data pathways and life-cycle. Since she is mostly responsible for data entry, she develops statistical models in order to ensure appropriate data quality. Therefore, the algorithms that she develops include descriptive statistics and comparative and regression analysis. She is interested in a platform where she will be able to create her own account and save her information.

USER STORIES

ID	User Story
DATA PROVIDER	
DP15	As a Senior Researcher in the education domain, I want to protect education data from misuse that aims in profiling students, so that I can defend individual data privacy
DP16	As a Senior Researcher in the education domain, I want to protect students from being automatically subjected to companies' laws and conditions that their institutions comply to, so that I can defend individual data privacy
DP17	As a Senior Researcher in the education domain, I want to provide access to my data without the risk of data loss or privacy, and without legal or ethical limitations, so that I can give data to energy and other domains that evaluate budget and consumption rates
DP18	As a Senior Researcher in the education domain, I want to provide access to my data in a secure and privacy aware way, but I do not want to exchange/send my data with others, so that I am sure that no misuse of my data will occur, and data will be maintained in my database
DP19	As a Senior Researcher in the education domain, I want to always keep control over the data to which I provide access, so that I am sure that no misuse of my data will occur

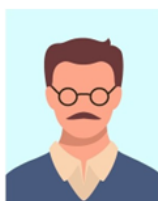
DP20	As a Senior Researcher in the education domain, I want to be able to decide the level of access I am willing to provide to the data I share with other people, so that I am sure that no misuse of my data will occur
CONSUMER	
C15	As a Researcher in the education domain, I want to be able to access and obtain data from economy, job market, energy, and health domains, so that I can extract useful insights for the advancement of the education domain
C16	As a Researcher in the education domain, I want to be able to easily confirm the authenticity of the data, so that I can use the best suitable datasets for my research
C17	As a Researcher in the education domain, I want to be able to easily access data from different data sources without restrictions, so that I can advance my research in the education domain
C18	As a Researcher in the education domain, I want to be able to perform computations and analysis over the data I access, so that I can extract useful insight about the education domain
C19	As a Researcher in the education domain, I want to know the way a website/portal collects data from various sources, so that I am sure that data privacy and protection from surveillance are respected regarding the data I want to access
C20	As a Researcher in the education domain, I want to be able to validate AI support decisions, so that I can use their outputs on specific interventions and evaluation of learning outcomes
C21	As a Researcher in the education domain, I want to utilize an AI model that could collect information from different websites/portals and track the use of data, so that I can further expand my research on how data are being used
MODEL PROVIDER	
MP15	As a Data Scientist in the education domain, I want to have access to support regarding cross platform functioning, so that I can utilize technical solutions easily without any expert's knowledge
MP16	As a Data Scientist in the education domain, I want to be able to register and create a new account, so that I can utilize the components that the platform may offer
MP17	As a Data Scientist in the education domain, I want to provide statistical models, so that I can assist researchers in other domains

MP18	As a Data Scientist in the education domain, I want to provide algorithms that include descriptive statistics and comparative and regression analysis, so that I can assist researchers in other domains

3.4 AUTOMOTIVE DOMAIN

PERSONAS

DATA PROVIDER



Jim Cross

Domain: Automotive
Age group: 50-60 years
Occupation: Technical Manager
Years of experience: 15+ years
Daily data usage: 2-4 hours
TRUSTEE Role: Data Provider

"I want to upload data easier and without manual effort"

Goals

- Data exchange with space and other domains
- Aggregate data from multiple vendors
- Collaboration with various domains, such as energy, transportation, climate, weather forecasting, and civil engineering
- Data synchronization while maintaining data integrity
- Data compression
- Share annotated non-personal data relating to sensory inputs, automotive simulators, and vehicles

Frustrations

- "It frustrates me that data annotation is of high demand, time-consuming, and requires expertise and manual effort"
- "It frustrates me that IPR and data privacy issues restrict data sharing"

Jim is a Technical Manager in the automotive domain mainly involved in software testing, integration, and review on data acquired by cameras, lidars, radars, and uss. Through his domain, he can provide data related to Federated Learning Output or Private Grounds captured video (fictitious), as well as Key Performance Indicators. Moreover, he can share data relating to photo-realistic artificially generated scenes, algos outputs, Key Performance Indicators, evaluation protocols, and validation benchmarks.

CONSUMER

"I want to access data for machine learning modules"



Clara Barton

Domain: Automotive
Age group: 40-50 years old
Occupation: Senior Researcher
Years of experience: 3-15 years
Daily data usage: 2-4 hours
TRUSTEE Role: Consumer

Goals

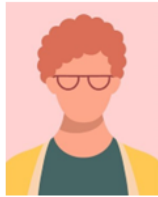
- Access to timestamped synchronized data
- Exchange data for secondary use with the space domain
- Aggregate data from multiple vendors and collaborate with various domains, such as energy, transportation, climate, weather forecasting, and civil engineering
- Access to text-based metadata generation for more efficient data querying
- Monitor and assessment of AI models regarding FAIR principles and security
- Reusability of algorithms
- Use the data I access for training AI models
- Access to indicators of runtime efficiency

Frustrations

- "It frustrates me that there is low quantity, quality, and diversity of available data"
- "It frustrates me that there are not enough data available from different geographic locations"
- "It frustrates me that data sharing requires big storage"
- "It frustrates me that IPR and data privacy issues restrict data sharing"
- "It frustrates me that obtained datasets often require pre-processing steps"
- "It frustrates me that collecting and aligning data from different domains is difficult"
- "It frustrates me that it often takes me a lot of time to understand how an AI model functions so I can trust its outputs"
- "It frustrates me that data modifications/computations require time, computation resources, and offline processing"
- "It frustrates me that high computational cost requires distributed processing"

Clara is a Senior Researcher in the automotive domain, and she is interested in accessing data from the space domain to assist her in the vehicle self-localization algorithms. She is also interested to have access to secondary use data from climate, weather forecasting, and transportation with city maps domains to combine them and create more realistic scenarios for simulations to enhance the performance of her developed algorithms. She is interested in AI and ML algorithms for object recognition, depth estimation, trajectory planning, fusion of Driver System Monitoring, and research on sensing function related to autonomous driving. She wants to use a platform where synergies among pillars, dataset generation, transfer learning, and reusability of algorithms is enabled. Furthermore, she would use a platform in which relevant data and datasets can be easily found.

MODEL PROVIDER



Patrick Kevinson

Domain: Automotive
Age group: 30-40 years
Occupation: Software Developer
Years of experience: 3-15 years
Daily data usage: 2-4 hours
TRUSTEE Role: Model Provider

"I want to modify data without computational costs"

Goals

- Generate artificial data based on real world diversity
- Training of ML models
- Development of federated modules
- Extension of extended AI models
- Transfer of learning
- Share annotated non-personal data relating to trained deep neural networks

Frustrations

- "It frustrates me that there is a lack of efficiency in data querying and data quality diagnostics tools"
- "It frustrates me that data synchronization is not yet achieved"
- "It frustrates me that research on non-text data is not efficient and needs to run on metadata"
- "It frustrates me that data analysis key performance indicators are not defined for the entire community"
- "It frustrates me that data cannot be queried because they usually exist in the form of files in folders"

Patrick is a Software developer mainly involved in the design and development of Machine Learning modules used for scene awareness, and innovative perception functions regarding Automotive Driving Assistance Systems and Autonomous Driving functions. His data usage activities may include, among others, sensor fusion, machine learning, sensor modelling or calibration, data parsing, visualization, and assessing the structural integrity of data by verifying that the signal quality and time synchronization comply to the specifications. His data processing often includes mathematical or logical operations, statistical processing, numerical optimization, machine learning, visualization, and data encoding or decoding. Furthermore, his data processing algorithms include optimization and training deep learning models algorithms, federated learning algorithms for SLAM tasks, and object detection tasks. He develops AI and ML algorithms relating to visual odometry, cross traffic alert, ego-vehicle self localization, co-operative localization, rear collision warning, road segmentation, simultaneous localization and parking or home-zone parking, valet parking, online sensor calibration, lane sensing and departure warning, three-dimensional environmental modelling, and trajectory planning. He would like to share his annotated non-personal data relating to trained deep neural networks, to train and evaluate machine/deep learning algorithms. He can provide various AI and deep learning tools for applying federated learning frameworks. He wants a platform in which relevant data can be easily found and training of machine learning modules used for automotive sensing are enabled. Furthermore, he wants the platform to enable transfer learning, synergies among pillars, training of machine learning modules used for automotive sensing, and dataset generation. Furthermore, he supports that datasets from various vendors could be gathered and utilized for AI and deep learning tools, thus, applying federated learning frameworks.

USER STORIES

ID	User Story

DATA PROVIDER	
DP2	As a Technical Manager in the automotive domain, I want to compress big data in an efficient and fast way, so that I can collaborate and share them with other domains, such as space, fast and without the need of big storage availability
DP3	As a Technical Manager in the automotive domain, I want to enable transfer of learning without IPR and data privacy issues, so that I can collaborate with other domains, such as space
DP4	As a Technical Manager in the automotive domain, I want to annotate data fast and without manual effort, so that I can share them with other domains, such as energy, transport, climate, weather forecasting, and civil engineering more efficiently
DP5	As a Technical Manager in the automotive domain, I want to be able to automatically synchronize datasets while keeping their integrity, so that I can update and enhance the knowledge I am sharing
DP6	As a Technical Manager in the automotive domain, I want to be certain that the data I provide will be used with respect to privacy regulations and ethical considerations, so that I am sure that I respect data privacy and legal regulations when providing access to my data
DP7	As a Technical Manager in the automotive domain, I want to be certain I will get informed regarding the way that my data are used (accessed, processed, utilized for the training of any model), so that I am always aware of how my data are used and I can terminate the procedure if any violation of my interests is happening
DP8	As a Technical Manager in the automotive domain, I want to share my annotated non-personal data relating to sensory inputs, automotive simulators, and vehicles, so that I can assist researchers in other domains
CONSUMER	
C2	As a Senior Researcher in the automotive domain, I want to have access to real-world, diverse, high-quality data, so that I can be able to generate artificial data based on real-world diversity
C3	As a Senior Researcher in the automotive domain, I want to process available, reliable, and diverse data without IPR and privacy issues, so that I can collaborate with other domains, such as space
C4	As a Senior Researcher in the automotive domain, I want to have access on text-based metadata, so that I can achieve efficient data querying
C5	As a Senior Researcher in the automotive domain, I want to access fully-formed, timestamped, and synchronized datasets, so that I can spend less time in their pre-processing

C6	As a Senior Researcher in the automotive domain, I want to access fully-formed, timestamped, and synchronized datasets, so that I can easily query and align them with other data from different domains
C7	As a Senior Researcher in the automotive domain, I want to be able to reuse already trained AI and ML algorithms, so that I can proceed with and advance my research faster and with more realistic scenarios
C8	As a Senior Researcher in the automotive domain, I want to be able to assess and monitor the performance of the AI models that I use regarding fairness, explainability, robustness, and security, so that I can trust their outputs
C9	As a Senior Researcher in the automotive domain, I want to utilize AI/ML algorithms for object recognition, depth estimation, trajectory planning, fusion of Driver System Monitoring, research on sensing function related to autonomous driving, and vehicle self-localization algorithms, so that I can further expand my research
C10	As a Senior Researcher in the automotive domain, I want to use the data that I access for feeding AI models, so that I can extract additional information about the data
C11	As a Senior Researcher in the automotive domain, I want to be able to understand and interpret how an AI model produced the additional knowledge that I wish to extract from the data that I access, so that I can modify and optimize my search queries and knowledge extraction requests, as well as trust the results that I receive
C12	As a Senior Researcher in the automotive domain, I want to perform data modifications more easily and, in a runtime, efficient manner, so that I do not have to invest much time in offline processing, and I do not constantly require expert's knowledge to do so
C13	As a Senior Researcher in the automotive domain, I want to perform data computations with low computational cost perform, so that I do not have to use distributed processing
C14	As a Senior Researcher in the automotive domain, I want to generate, process, and modify big data in an efficient and fast way, so that the constant need of domain expert's knowledge is omitted
MODEL PROVIDER	
MP2	As a Software Developer in the automotive domain, I want to be able to generate artificial data based on real-world diversity, so that I can develop and train ML and AI models and federated modules, and even extend current AI models
MP3	As a Software Developer in the automotive domain, I want to use data quality diagnostic tools, so that I can find the most suitable data for the development of federated modules and the extension of extended AI models

MP4	As a Software Developer in the automotive domain, I want to process data fast and easily with defined data analysis key performance indicators, so that I can use them more efficiently in my developed algorithms
MP5	As a Software Developer in the automotive domain, I want to share my annotated non-personal data relating to trained deep neural networks, so that I can assist in training and evaluation of machine/deep learning algorithms
MP6	As a Software Developer in the automotive domain, I want to provide AI and deep learning tools for data, so that I can apply federated learning frameworks
MP7	As a Software Developer in the automotive domain, I want to provide AI/ML algorithms for scene awareness, automotive sensing, optimization and training deep learning models, federated learning for SLAM tasks, object detection tasks, visual odometry, cross traffic alert, ego-vehicle self localization, co-operative localization, rear collision warning, road segmentation, simultaneous localization and parking or home-zone parking, valet parking, online sensor calibration, lane sensing and departure warning, three-dimensional environmental modelling, trajectory planning, and innovative perception functions regarding Automotive Driving Assistance Systems and Autonomous Driving functions, so that I can assist researchers in other domains

3.5 SPACE DOMAIN

PERSONAS

DATA PROVIDER



Samantha Collier

Domain: Space
Age group: 60+ years
Occupation: Senior Researcher
Years of experience: 15+ years
Daily data usage: 2-4 hours
TRUSTEE Role: Data Provider

"I want to share my processed data and combine them with data from other domains"

Goals

- Give statistics data for secondary use to seismologists, key infrastructure researchers, or others for research
- Give data for secondary use to healthcare, education, medical, and sport performance related domains
- Share data catalogues, tutorials, and extraction or analysis algorithms
- Protection of personal data and informed consent

Frustrations

- "It frustrates me that sharing data is difficult due to privacy policies"
- "It frustrates me that I have to apply TLS encryption or anonymization procedures in order to share my data"

Samantha is a Senior Researcher in the space domain. She shares publicly available data either by referring to the source or by modifying them and sharing them with her colleagues internally. She ensures data quality by engaging qualified personnel in the fields of interest. She can provide name, identification number, physical, psychological, and genetic data. The information she extracts from the publicly available data can be relevant to a broader community, such as seismologists or key infrastructure researchers, that utilize statistics from the space domain. She ensures data quality based on the specialization knowledge and the expertise of her work field colleagues. She believes that the business medical sector can use the space data to improve the evaluation of patients. She would be interested in a platform where she would be able to provide data and perform simple data processing in a way that the data itself will not be visible to its users.

CONSUMER



Kelvin Grey

Domain: Space
Age group: 50-60 years
Occupation: Project Manager
Years of experience: 15+ years
Daily data usage: 2-4 hours
TRUSTEE Role: Consumer

"I want to access data and combine them with other domains"

Goals

- Get and combine data for secondary use from health, medical, energy, agriculture, behavioural, and manufacturing domains
- Access to large volumes of healthcare data
- Perform data manipulation on privacy-preserved data
- Valid data provided by authenticated users
- Advancement of science and increase of societal impact

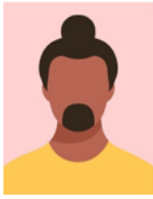
Frustrations

- "It frustrates me that historical data often set on different standards, are less interoperable and require lots of modifications"
- "It frustrates me that various datasets, such as power outage data for smaller incidents, are not publicly available"
- "It frustrates me that sharing data is difficult due to privacy policies"
- "It frustrates me that most of data are correlated with personal data making them difficult to share, even with pseudonymization procedures"
- "It frustrates me that centralized databases lack clinical data"
- "It frustrates me that special data querying takes time to generate results and correlation from multiple repositories is difficult"
- "It frustrates me that large volume datasets require specialized infrastructures, especially with near real-time processing, and are often flawed, forcing one to work with subsets"
- "It frustrates me that there are non-cross-platform results and many automatization capabilities in data manipulation"
- "It frustrates me that I have to learn specific program commands for personal input methods in order to modify data"

Kelvin is a Project Manager in the space domain. His data usage activities relate to conducting numerical simulation based on pools of data, extracting data from space catalogues, editing and filtering these data, graphical and scientific interpretation of resulted data, and publishing catalogues. He is interested to process data from the healthcare, energy, and agriculture domains, as well as combine them with data from the space domain. He also wants data from the clinical domain to research relations to the development of countermeasures and access large volume of data processing for digital profiling solutions. He is interested in AI/ML algorithms for data extraction and update, automatization of running simulations, output and filtering of resulted data, data analysis based on previous work results, and evaluation of psychophysiological status of the aerospace personnel involved in long-term space missions. He wants to use a platform with a variety of manipulated data, and registered users, as well as be able to perform fast simulations and study less-complex scenarios. Moreover, he wants to utilize valid data that are provided by authenticated users. Furthermore, he wants the platform to support integration that can solve many existing problems with undesirable effects in the data sharing between sectors.

MODEL PROVIDER

"I want a platform where I can manipulate a variety of different data"



Travis Santos

Domain: Space
Age group: 30 – 40 years
Occupation: Data Scientist
Years of experience: 3 – 15 years
Daily data usage: 2 – 4 hours
TRUSTEE Role: Model Provider

Goals

- Automate repetitive processes
- Utilize knowledge produced by other experts in the fields
- Eliminate time consumption in processes that have already been conducted in the past by others

Frustrations

- "It frustrates me that I have to use different environments for interpretation of the data based on my requirements"
- "It frustrates me that there are filtering limitations due to the range of data parameters in the database"
- "It frustrates me that I have limited tool options and expensive proprietary software in data analysis"
- "It frustrates me that data analysis is heavily influenced by personal knowledge and expertise of reachable specialists in the field"
- "It frustrates me that it is difficult to personalize the code to an activity and generate higher number of training data"
- "It frustrates me that, regarding data entry, there are mismatches in data structuring between different platforms"

Travis is a Data Scientist in the space domain. His data usage activities relate to, among others, conducting numerical simulations, extracting, editing, filtering data from space catalogues, and graphical and scientific interpretation of resulted data. The computations he usually performs on data include basic and complex mathematical operations and graphic outputs or lists of tables. He produces several variants of the datasets that he usually works with to simplify the data processing cycle. The data analysis that he usually performs either includes advanced 3D visualization, or processing tools that extract specific information from the datasets. Although the processing tools he utilizes perform simple mathematical operations, the value they extract can be a rather large number. The algorithms he deploys include computational statistics algorithms, computer vision, filtering and categorizing data, exporting to different filetypes and different outputs, and digital signal, image, and audio processing. He would be interested in a platform where he would be able to provide data and perform simple data processing in a way that the data itself will not be visible to its users. He also wants a platform with a variety of manipulated data, and registered users, as well as be able to perform fast simulations and study less-complex scenarios. Furthermore, he would like this platform to resemble well-known data repositories with the enhancement of the availability of code maintenance and communication with the code creators.

USER STORIES

ID	User Story
DATA PROVIDER	
DP71	As a Senior Researcher in the space domain, I want to share statistics data for secondary use with seismologists, key infrastructure researchers, etc., so that I can provide additional knowledge and assist in advancing scientific research
DP72	As a Senior Researcher in the space domain, I want to overcome data privacy policies restrictions, so that I can provide access to my data for secondary use to the education domain
DP73	As a Senior Researcher in the space domain, I want to overcome data privacy policies restrictions, so that I can share data catalogs, tutorials, and analysis outputs easier
DP74	As a Senior Researcher in the space domain, I want to automatically encrypt and secure my data, so that I can provide access to my data to healthcare, education, medical, and sport performance related domains
DP75	As a Senior Researcher in the space domain, I want to protection of personal data and informed consent, so that I can securely exchange data with other domains
CONSUMER	

C71	As a Project Manager in the space domain, I want to have access to more data, such as power outage data for smaller incidents, or data from health, energy, and agriculture domains, so that I can perform data manipulation on privacy-preserved data
C72	As a Project Manager in the space domain, I want to access historical data that are set on the same standards, are interoperable, and don't require modifications, so that I can perform queries and analyse them faster and more efficiently
C73	As a Project Manager in the space domain, I want to overcome data privacy policies restrictions, so that I can be granted access to secondary use data from the manufacturing domain
C74	As a Project Manager in the space domain, I want to eliminate the correlation of data with personal data, so that I can receive medical, healthcare, and behavioral data
C75	As a Project Manager in the space domain, I want to access clinical data, so that I can research relations to the development of countermeasures and access large volume of data processing for digital profiling solutions
C76	As a Project Manager in the space domain, I want to use centralized databases that contain clinical and healthcare data, so that I can research relations to the development of countermeasures and access large volume of data processing for digital profiling solutions
C77	As a Project Manager in the space domain, I want to automatically correlate and query data from multiple sources, so that I can produce better results in my data analysis
C78	As a Project Manager in the space domain, I want to access valid data that are provided by authenticated users, so that I can produce better results in my data analysis
C79	As a Project Manager in the space domain, I want to utilize AI/ML algorithms for data extraction and update, automatization of running simulations, output and filtering of resulted data, data analysis based on previous work results, and evaluation of psychophysiological status of the aerospace personnel involved in long-term space missions, so that I can further expand my research
C80	As a Project Manager in the space domain, I want to use generalized program commands for data modification, so that I can modify data easier and without the need of expert's knowledge
C81	As a Project Manager in the space domain, I want to manipulate all data in one environment that enables automatization techniques and produce cross-platform results, so that I can faster and more efficiently utilize the results to extend my research and knowledge expertise
C82	As a Project Manager in the space domain, I want to process valid large volume datasets without the need for specialized infrastructures, so that I can produce results faster, which can assist in the advancement of science and increase of societal impact

MODEL PROVIDER	
MP71	As a Data Scientist in the space domain, I want to be able to personalize the code for training data generation easier, so that I can create a greater and more suitable amount of training data faster
MP72	As a Data Scientist in the space domain, I want to have more data parameter options/fields in a database, so that I can later have more filtering options and perform special data querying faster
MP73	As a Data Scientist in the space domain, I want to use a data repository where code can be easily maintained and where I can communicate faster with code creators, so that I can edit and publish more of my scientific research, as well as expand my expertise and knowledge
MP74	As a Data Scientist in the space domain, I want to be able to reach and communicate with more specialists in a specific field, so that I can gain insight in the field and perform better data analysis
MP75	As a Data Scientist in the space domain, I want to use a platform that stores knowledge extracted from processes that other users have conducted in the past, so that I can get results faster in case the process I want to conduct has already been conducted by someone else before
MP76	As a Data Scientist in the space domain, I want to use a platform that offers the ability for automating procedures regarding data consumption and other data processes, so that I can replicate similar procedures and eliminate time consumption due to repetitive tasks to focus on my research
MP77	As a Data Scientist in the space domain, I want to have more tool options and free software for data analysis, so that I can produce better results faster and easier
MP78	As a Data Scientist in the space domain, I want to provide AI/ML algorithms relating to computational statistics, computer vision, filtering and categorizing of data, exporting to different filetypes and different outputs, and digital signal, image, and audio processing, so that I can assist researchers in other domains

3.6 MULTI-DISCIPLINARY DATA EXCHANGE

PERSONAS

DATA PROVIDER



Erika Daves

Domain: Multi-Disciplinary
Age group: 30-40 years
Occupation: Senior Researcher
Years of experience: 3-15 years
Daily data usage: <1 hour
TRUSTEE Role: Data Provider

"I want to be able to share my data securely"

Goals

- Secure exchange of data
- Allow access of data only to authenticated users

Frustrations

- "It frustrates me that privacy regulations, legal frameworks, lack of trust and defective interoperability make it difficult to exchange data for secondary use"

Erika is a Data Provider in the multi-disciplinary domain and her typical day at work includes contributing to different projects either as a member or a manager. Her data usage processes may include data entry, manipulation, modification, querying, sharing, and analysis. She can provide non-personal and personal consumption data to other domains. She is interested in a platform that allows the provided data to be accessed only by authenticated users.

CONSUMER



Eric Holland

Domain: Multi-Disciplinary
Age group: 40-50 years
Occupation: Project Manager
Years of experience: 3-15 years
Daily data usage: 2-4 hours
TRUSTEE Role: Consumer

"I want to access the exchange of data between different domains"

Goals

- Get data for secondary use from climate and weather domains
- Resolve data exchange technical challenges between heterogeneous systems and data repositories
- Exchange data with energy, healthcare, automotive, mobility, transport, and finance domains
- Give data for secondary use to the manufacturing domain
- Research data exchange between transport and energy domains
- Inherit AI operations from other domains and migrate their relevant AI/ML processes

Frustrations

- "It frustrates me that data repositories are not open-accessed"
- "It frustrates me that data and metadata are often inadequate, incomplete, or of bad quality"
- "It frustrates me that privacy regulations, legal frameworks, lack of trust, and defective interoperability make it difficult to exchange data for secondary use"
- "It frustrates me that I often have to develop or modify the data transformation for each use case, due to data format inconsistencies between different companies"
- "It frustrates me that data querying is done with SPARQL, which is a non-user-friendly, complex, and error-prone language"

Eric is a Data Consumer in the multi-disciplinary domain. He is interested in many data domains and particular in identifying and extracting knowledge from the integration/exchange of data that is collected from different domains. He believes that the energy, healthcare, and automotive domain could benefit from exchanging data with each other, and he would like to access their data for secondary use. Moreover, he is interested in the integration of data between the space and energy domains, and health with education domains. Furthermore, he wants to learn how to address the technical challenges set by data exchange between heterogeneous systems and data repositories. He would be interested in AI and ML algorithms for prediction, automation, and image object detection models. He is interested in a platform where he can access different data spaces and domains that could create different multi-disciplinary scenarios. He also wants to utilize a platform that has clear usage rules, includes high quality data and metadata, has enabled traceability of potential cleansing operations, has easy access and permissions, and has unbiased basic services operators. He considers that it would be important for such a platform to comply with the requirements of data sovereignty, neutrality, traceability, and data interoperability and findability. He would be interested in utilizing a platform that has clear usage rules, includes high quality data and metadata, has enabled traceability of potential cleansing operations, has easy access and permissions, and has unbiased basic services operators. Furthermore, in that platform, he would like to be able to utilize and apply common semantic ontologies to support the harmonization of datasets. Also in that platform, he would like to access heterogeneous data and be able to utilize and apply common semantic ontologies to support their harmonization.

MODEL PROVIDER



Nora Austin

Domain: Multi-Disciplinary

Age group: 30-40 years

Occupation: Data Scientist

Years of experience: 3-15 years

Daily data usage: 2-4 hours

TRUSTEE Role: Model Provider

"I want utilize data from different domains to develop new models"

Goals

- Identify and extract knowledge from the integration and exchange of collected data from different domains, such as space with energy, health with education

Frustrations

- "It frustrates me that most of the times there are no examples or tutorials on how to use a new solution/platform"

Nora is a TRUSTEE Developer in the multi-disciplinary domain and her typical day at work includes attending project meetings and software development for these projects. Her data usage processes may include data entry, manipulation, modification, querying, sharing, and analysis. Depending on the task at hand, the computations he usually performs include visualization and mathematical and logical operations relating, among others, to statistical analysis, distributions, variances, outliers, and normalizations. She would also be interested in inheriting AI operations from other domains and migrate their relevant AI/ML processes. She wants to access data from many domains and particular to identify and extract knowledge from the integration/exchange of data that is collected from these different domains, such as the integration of data between the space and energy domains, and health with education domains. Furthermore, she is interested in learning how to address the technical challenges set by data exchange between heterogeneous systems and data repositories. Furthermore, she would like to access heterogeneous data through this platform and be able to utilize and apply common semantic ontologies to support their harmonization, as well as research on the different cross-sectoral use cases.

USER STORIES

ID	User Story
DATA PROVIDER	
DP62	As a Senior Researcher in the multi-disciplinary domain, I want to use a secure platform that respects privacy regulations, so that I can share data easier with trust and effective interoperability
DP63	As a Senior Researcher in the multi-disciplinary domain, I want to use a platform where only authenticated users can have access to the provided data, so that I can ensure security and use of secondary use data
CONSUMER	
C62	As a Project Manager in the multi-disciplinary domain, I want to use a platform that resolves the format inconsistencies between data coming from different companies and heterogeneous systems or repositories, so that I can work with data without the exchange technical challenges
C63	As a Project Manager in the multi-disciplinary domain, I want to perform data querying through user-friendly, simple, and error-free programming languages, so that I can find the information I am looking for faster and more efficiently

C64	As a Project Manager in the multi-disciplinary domain, I want to access data easier while complying to the privacy regulations and legal frameworks, so that I can exchange data with energy, healthcare, automotive, mobility, transport, and finance domains, and give data to the manufacturing domain
C65	As a Project Manager in the multi-disciplinary domain, I want to use more open-accessed data repositories, so that I can utilize secondary use data from the climate and weather domains
C66	As a Project Manager in the multi-disciplinary domain, I want to have access to adequate, complete, and good quality data and metadata, so that I can effectively research the data exchange between transport and energy domains, and other different cross-sectoral use cases
C67	As a Project Manager in the multi-disciplinary domain, I want to inherit AI/ML operations from other domains, so that I can migrate their relevant AI/ML processes
C68	As a Project Manager in the multi-disciplinary domain, I want to utilize AI and ML algorithms for prediction, automation, and image object detection models, so that I can further expand my research
C69	As a Project Manager in the multi-disciplinary domain, I want to be able to utilize and apply common semantic ontologies, so that I can support the harmonization of datasets
C70	As a Project Manager in the multi-disciplinary domain, I want to utilize a platform that has clear usage rules, enabled traceability of potential cleansing operations, easy access and permissions, and unbiased basic services operators, so that I can assist my research in the multi-disciplinary domain
MODEL PROVIDER	
MP62	As a Data Scientist in the multi-disciplinary domain, I want to study the integration and exchange of data between space and energy, as well as health and education, so that I can extract additional knowledge
MP63	As a Data Scientist in the multi-disciplinary domain, I want to be able to apply common semantic ontologies to heterogeneous data, so that I can support their harmonization
MP64	As a Data Scientist in the multi-disciplinary domain, I want to have access to examples or tutorials on how to use a new solution or platform, so that I can use that platform or solution effectively

ALL DOMAINS

PERSONAS

DEVELOPER



Larry Moran

Age group: <30 years

Occupation: Software Developer

Years of experience: 3 - 15 years

Daily data usage: 2 - 4 hours

TRUSTEE Role: Developer

"I want to be able utilize APIs of platforms so that I can make use of their offered functionalities"

Goals

- Utilize potential extensions for SDK integration in other tools
- Utilize a platform that contains APIs or portals
- Smoothly working APIs
- Access AI/ML code packages or tutorials to extend personal studies and optimize working simulations
- Improved inbuilt visualization tools

Frustrations

- "It frustrates me that platforms don't include good APIs and proper documentation"

Larry is a Software Developer mainly involved in the design and development of Machine Learning components that can be used for scene awareness, as well as innovative perception functions regarding Automotive Driving Assistance Systems and Autonomous Driving functions. His data usage processes involve collecting and monitoring of 5G infrastructure and I4.0 testbed data, data querying through the monitoring data collection, and data sharing via the IDS data exchange between testbeds. His data usage activities may include, among others, sensor fusion, machine learning, sensor modelling or calibration, and visualization. His data processing often includes mathematical or logical operations, statistical processing, numerical optimization, machine learning, visualization, and data encoding or decoding. He would like to utilize a platform that is compatible with existing data exchange technology infrastructures, such as IDS and Gaia-X. He believes that IDS can assist the space domain in generating a relevant connector to exchange data with other domains. Moreover, he would like to utilize potential extensions for SDK integration in other tools and have example applications and tutorials for each use case usage. He would be interested in a platform that contains good and valid APIs with proper documentation. He would also like to be able to deploy different IDS connectors for different providers for further research. Moreover, he would like to be able to create different multi-disciplinary scenarios.

USER STORIES

ID	Description
DEVELOPER	
D83	As a Software Developer, I want to utilize a platform that contains good and valid APIs with proper documentation, so that I can perform requests to expand my research
D84	As a Software Developer, I want to utilize extensions for SDK integration in other tools, so that I can exploit the offered solution as much as possible and advance my research
D85	As a Software Developer, I want to have access to datasets from the health domain, so that I can deploy different IDS connectors for different healthcare providers
D86	As a Software Developer, I want to use a user-friendly platform that will be open, free, and contain good inbuilt visualization tools and APIs, so that I can further expand my research
D87	As a Software Developer, I want to utilize IDS connectors, so that I can generate a relevant connector to exchange data with other domains
D88	As a Software Developer, I want to use a platform that is compatible with existing data exchange technology infrastructures, such as IDS and Gaia-X, so that I can further expand my research

D89	As a Software Developer, I want to access different APIs or portals, so that I can create different multi-disciplinary scenarios
D90	As a Software Developer, I want to utilize potential extensions for SDK integration in other tools, and have example applications and tutorials for each use case usage, so that I can further expand my research
D91	As a Software Developer, I want to access AI/ML code packages or tutorials, so that I can extend personal studies and optimize working simulations

ANNEX 4: USER-CENTRED REQUIREMENTS AND INTERCONNECTION WITH LEGAL REQUIREMENTS

Requirement ID	Personas (Data Provider, Moder Provider, Consumer, Developer)	User Story ID (from which User Story the requirement stems out)	Requirement (stemming from User Stories)	Priority (M: Must-have. Mandatory Requirement, S: Should-have. Desirable Requirement, C: Could-have. Optional Requirement, W: Will-not-have. Possible Future Enhancement)	Interconnected Legal requirements from table inserted in Annex “TRUSTEE Legal Requirements”	Corresponding legal requirements for users	Architecture Building Block (that could address the requirement)
USR-001	Data Provider, Consumer, Developer	DP46, DP49, DP71, C3, C10, C15, C23, C31, C32, C33, C34, C35, C39, C45, C46, C47, C48, C51, C53, C54, C64, C66, C71, C75, D86	TRUSTEE must enable collaboration among domains and enable access and use of data for secondary use from multiple domains	Mandatory Requirement	LEG-TSD-013 LEG-TSD-014 LEG-TSD-036, LEG-TSD-037-039, LEG-TSD-051-062, LEG-TSD-064, LEG-TSD-066, LEG-TSD-068-071, LEG-TSD-074-080, LEG-TSD-082, LEG-TSD-085-088, LEG-TSD-103, LEG-TSD-111-113.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-066, LEG-USR-068-069 LEG-USR-070-072, LEG-USR-079-080, LEG-USR-083, LEG-USR-091, LEG-USR-093-099.	DL, HEDF, ATR, SSI-HE, STM, DPIA
USR-002	Data Provider	DP24	TRUSTEE must use standardized methods and clear instructions for correct data entry and formatting	Mandatory Requirement	LEG-TSD-005, LEG-TSD-006, LEG-TSD-008, LEG-TSD-014, LEG-TSD-017, LEG-TSD-018, LEG-TSD-020,	LEG-USR-055, LEG-USR-068, LEG-USR-070, LEG-USR-079, LEG-USR-083, LEG-UST-086,	DL, SSI-HE, HEDF

					LEG-TSD-021-022, LEG-TSD-026, LEG-TSD-041, LEG-TSD-053-062, LEG-TSD-069, LEG-TSD-071, LEG-TSD-074, LEG-TSD-080, LEG-TSD-082, LEG-TSD-083, LEG-TSD-085, LEG-TSD-095, LEG-TSD-097, LEG-TSD-098, LEG-TSD-099, LEG-TSD-100, LEG-TSD-103, LEG-TSD-108.	LEG-USR-091, LEG-USR-094.	
USR-003	Data Provider	DP18	Users must provide access to their data in a secure and privacy aware way without sending their raw data to others	Mandatory Requirement	LEG-TSD-003-005, LEG-TSD-012-037, LEG-TSD-039, LEG-TSD-042-050, LEG-TSD-051-062, LEG-TSD-066-068, LEG-TSD-070-073, LEG-TSD-075-080, LEG-TSD-084-091, LEG-TSD-104, LEG-TSD-108-113.	LEG-USR-001-003, LEG-USR-007-008, LEG-USR-010, LEG-USR-012-025, LEG-USR-027-029, LEG-USR-031-032, LEG-USR-035, and related LEG-USR-036-050, LEG-USR-055-061, LEG-USR-063-066, LEG-USR-068-070, LEG-USR-072-073, LEG-USR-075, LEG-USR-076-078, LEG-USR-079-080, LEG-USR-094, LEG-USR-095, LEG-USR-099.	DL, HEDF, AM, DPIA, STM
USR-004	Data Provider, Consumer	DP52, C70	TRUSTEE must support more referential integrity in data modification	Mandatory Requirement	LEG-TSD-004, LEG-TSD-013, LEG-TSD-014, LEG-TSD-021, LEG-TSD-026, LEG-TSD-069, LEG_TSD-085, LEG-TSD-095, LEG-TSD-109, LEG-TSD-110.	LEG-USR-002-003, LEG-USR-006, LEG-USR-015-016, LEG-USR-018-021, LEG-USR-028, LEG-USR-029, LEG-USR-035-050, LEG-USR-052-053, LEG-USR-055-056, LEG-USR-058-059,	ATR

						LEG-USR-061, LEG-USR-065-068, LEG-USR-070-072, LEG-USR-074, LEG-USR-079, LEG-USR-083, LEG-USR-095.	
USR-005	Data Provider, Consumer	DP6, DP17, DP62, C17, C64, C73	The data provided must be used with respect to privacy regulations, legal frameworks, and ethical considerations and without the risk of data loss	Mandatory Requirement	LEG-TSD-001-117.	LEG-USR-001-099.	AM, ATR, DPIA, STM, TAI-SDF, HEDF
USR-006	Data Provider	DP7	The user must always be aware of how their data are used and they must be able to terminate the procedure if any violation of their interests may happen	Mandatory Requirement	LEG-TSD-001-003, LEG-TSD-005, LEG-TSD-011-038, LEG-TSD-055-58, LEG-TSD-066-067, LEG-TSD-070-091, LEG-TSD-098, LEG-TSD-102-103, LEG-TSD-106.	LEG-USR-001-050, LEG-USR-054, LEG-USR-057, LEG-USR-058-059, LEG-USR-061, LEG-USR-063, LEG-USR-065-066, LEG-USR-068-069, LEG-USR-070, LEG-USR-072-074, LEG-USR-080, LEG-USR-086, LEG-USR-094-099.	DPIA, STM, ATR
USR-007	Data Provider	DP19	Users must always keep control over the data to which they provide access	Mandatory Requirement	LEG-TSD-001-003, LEG-TSD-005, LEG-TSD-011-037, LEG-TSD-055-058, LEG-TSD-066-067, LEG-TSD-069-091, LEG-TSD-098, LEG-TSD-102-103, LEG-TSD-106, LEG-TSD-111-113.	LEG-USR-001-052, LEG-USR-054, LEG-USR-057, LEG-USR-058-059, LEG-USR-061, LEG-USR-063, LEG-USR-065-066, LEG-USR-068-069, LEG-USR-070, LEG-USR-072-074, LEG-USR-080, LEG-USR-086, LEG-USR-094-099.	DPIA, STM
USR-008	Data Provider	DP63	Only authenticated users must have access to the data of the providers	Mandatory Requirement	LEG-TSD-001-003, LEG-TSD-011-037, LEG-TSD-055-058, LEG-TSD-068-069,	LEG-USR-001-052, LEG-USR-054, LEG-USR-057, LEG-USR-058-059,	AM

					LEG-TSD-072, LEG-TSD-075-078, LEG-TSD-089, LEG-TSD-109-113.	LEG-USR-061, LEG-USR-063, LEG-USR-065-066, LEG-USR-068-069, LEG-USR-070, LEG-USR-072-074, LEG-USR-080, LEG-USR-086, LEG-USR-094-099.	
USR-009	Data Provider	DP15	TRUSTEE must protect education data from misuse that aims in profiling students	Mandatory Requirement	LEG-TSD-001-006 LEG-TSD-012 LEG-TSD-014-015 LEG-TSD-018-019 LEG-TSD-037 LEG-TSD-039 LEG-TSD-042-050 LEG-TSD-066 LEG-TSD-068-079 LEG-TSD-084 LEG-TSD-089-091 LEG-TSD-092-103.	LEG-USR-001-037, LEG-USR-041-042, LEG-USR-045-050, LEG-USR-057-063, LEG-USR-065-069, LEG-USR-070-072, LEG-USR-074, LEG-USR-083-092, LEG-USR-095, LEG-USR-099.	DPIA, STM, HEDF
USR-010	Data Provider	DP16	TRUSTEE must protect students from being automatically subjected to companies' laws and conditions	Mandatory Requirement	LEG-TSD-001-006 LEG-TSD-012 LEG-TSD-014-015 LEG-TSD-018-019 LEG-TSD-037 LEG-TSD-039 LEG-TSD-042-050 LEG-TSD-066 LEG-TSD-068-079 LEG-TSD-084 LEG-TSD-089-091 LEG-TSD-092-103.	LEG-USR-001-037, LEG-USR-041-042, LEG-USR-045-050, LEG-USR-057-063, LEG-USR-065-069, LEG-USR-070-072, LEG-USR-074, LEG-USR-083-092, LEG-USR-095, LEG-USR-099.	DPIA, STM
USR-011	Model Provider	MP16	TRUSTEE must enable registration and creation of a new account for each user	Mandatory Requirement	LEG-TSD-012 LEG-TSD-013 LEG-TSD-016 LEG-TSD-017-022	LEG-USR-001-037, LEG-USR-041-042, LEG-USR-045-050, LEG-USR-057-063,	AM

					LEG-TSD-025 LEG-TSD-038 LEG-TSD-043-050 LEG-TSD-051-062 LEG-TSD-079 LEG-TSD-085 LEG-TSD-087-089.	LEG-USR-065-069, LEG-USR-070-072, LEG-USR-074, LEG-USR-083-092, LEG-USR-095, LEG-USR-099.	
USR-012	Data Provider	DP47	TRUSTEE should accept data in different formats with unequivocal labels	Desirable Requirement	LEG-TSD-003, LEG-TSD-005, LEG-TSD-011-014, LEG-TSD-020, LEG-TSD-026, LEG-TSD-056, LEG-TSD-069, LEG-TSD-074, LEG-TSD-082-83, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-003, LEG-USR-015-016, LEG-USR-018, LEG-USR-026, LEG-USR-055-056, LEG-USR-058-060, LEG-USR-068, LEG-USR-079, LEG-USR-080, LEG-USR-083.	DL, HEDF
USR-013	Data Provider	DP24	TRUSTEE should make data provider's data readily available and without the need for them to reformat them	Desirable Requirement	LEG-TSD-008, LEG-TSD-013-016, LEG-TSD-038, LEG-TSD-041, LEG-TSD-055-056, LEG-TSD-069, LEG-TSD-073-074, LEG-TSD-077, LEG-TSD-082-083, LEG-TSD-085, LEG-TSD-088.	LEG-USR-003, LEG-USR-005, LEG-USR-015-016, LEG-USR-018, LEG-USR-026-029, LEG-USR-055, LEG-USR-058-059, LEG-USR-061, LEG-USR-062, LEG-USR-066, LEG-USR-068-069, LEG-USR-071-073, LEG-USR-076, LEG-USR-078-070, LEG-USR-083, LEG-USR-091, LEG-USR-093, LEG-USR-097-099.	DL, SSI-HE, HEDF
USR-014	Developer	D89	TRUSTEE should offer APIs or portals for some of its functionalities to be used	Desirable Requirement	LEG-TSD-026-027, LEG-TSD-032-034, LEG-TSD-111-113.	LEG-USR-003-053, LEG-USR-079-080, LEG-USR-081-093, LEG-USR-094, LEG-USR-095.	OneSS (SDK connectors), Dashboard

USR-015	Consumer	C65	TRUSTEE should be an open-accessed data repository	Desirable Requirement	LEG-TSD-012-014, LEG-TSD-067, LEG-TSD-069-070, LEG-TSD-075-078, LEG-TSD-081-085, LEG-TSD-111-113.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-094-099.	KR
USR-016	Consumer	C22, C72	TRUSTEE should enable interoperability and integration of data from different data sources	Desirable Requirement	LEG-TSD-013-014, LEG-TSD-069, LEG-TSD-074, LEG-TSD-081, LEG-TSD-083, LEG-TSD-085, LEG-TSD-112-113.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-083, LEG-USR-086, LEG-USR-088, LEG-USR-091, LEG-USR-096-099.	DL, HEDF, SSI-HE
USR-017	Consumer	C25	TRUSTEE should be a standardized interoperability platform	Desirable Requirement	LEG-TSD-069, LEG-TSD-074, LEG-TSD-083, LEG-TSD-085.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-083, LEG-USR-086, LEG-USR-088, LEG-USR-091, LEG-USR-096-099.	OneSS (SDK Toolbox), DL

USR-018	Developer	D87, D88	TRUSTEE should be compatible with existing data exchange technology infrastructures, such as IDS and Gaia-X	Desirable Requirement	LEG-TSD-069, LEG-TSD-075, LEG-TSD-081, LEG-TSD-083, LEG-TSD-085.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-083, LEG-USR-086, LEG-USR-088, LEG-USR-091, LEG-USR-096-099.	HEDF, DL, External Connectors
USR-019	Consumer	C16	TRUSTEE should confirm the authenticity of the data in an easy way	Desirable Requirement	LEG-TSD-013-016, LEG-TSD-022 LEG-TSD-085-086, LEG-TSD-095, LEG-TSD-099, LEG-TSD-103.	LEG-USR-001, LEG-USR-015, LEG-USR-026, LEG-USR-059, LEG-USR-065-066, LEG-USR-068, LEG-USR-073, LEG-USR-079-080, LEG-USR-083, LEG-USR-091, LEG-USR-094, LEG-USR-096 LEG-USR-098.	AM, DL, STM
USR-020	Consumer	C56	The user should be able to find out about the validity, quality, and fairness of the data	Desirable Requirement	LEG-TSD-013-016, LEG-TSD-026, LEG-TSD-033, LEG-TSD-056, LEG-TSD-058, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001-002, LEG-USR-007, LEG-USR-014, LEG-USR-015-016, LEG-USR-018, LEG-USR-019-020, LEG-USR-024, LEG-USR-026, LEG-USR-029, LEG-USR-034-050, LEG-USR-052, LEG-USR-055, LEG-USR-058, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-083, LEG-USR-086, LEG-USR-088-093.	KR, SSI-HE, DL, STM

USR-021	Consumer	C2, C3	TRUSTEE will not include diverse data	Desirable Requirement	LEG-TSD-069, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001, LEG-USR-015, LEG-USR-026, LEG-USR-055-056, LEG-USR-058, LEG-USR-068, LEG-USR-079, LEG-USR-083, LEG-USR-086, LEG-USR-091.	DL
USR-022	Consumer	C5, C6	TRUSTEE should enable access to fully-formed, timestamped, and synchronized datasets	Desirable Requirement	LEG-TSD-012-014, LEG-TSD-018, LEG-TSD-021-022, LEG-TSD-069, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103, LEG-TSD-109-110.	LEG-USR-001, LEG-USR-015, LEG-USR-025-026, LEG-USR-055-056, LEG-USR-058, LEG-USR-068, LEG-USR-079, LEG-USR-083, LEG-USR-086, LEG-USR-091.	KR, DL, ATR
USR-023	Consumer	C49	TRUSTEE should enable access to numerical and descriptive data	Desirable Requirement	LEG-TSD-018, LEG-TSD-069, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001, LEG-USR-024-026, LEG-USR-055-056, LEG-USR-058, LEG-USR-068, LEG-USR-079, LEG-USR-083, LEG-USR-086, LEG-USR-091.	DL
USR-024	Consumer	C55	TRUSTEE should include data with continuous distribution	Desirable Requirement	LEG-TSD-018, LEG-TSD-043-48, LEG-TSD-069, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001, LEG-USR-015-016, LEG-USR-018, LEG-USR-024-027, LEG-USR-030, LEG-USR-033-035, LEG-USR-045-050, LEG-USR-055-056, LEG-USR-058, LEG-USR-068, LEG-USR-079, LEG-USR-083,	

						LEG-USR-088, LEG-USR-091.	
USR-025	Consumer	C24	TRUSTEE should include higher volumes of data	Desirable Requirement	LEG-TSD-069, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001, LEG-USR-015-016, LEG-USR-018, LEG-USR-024-027, LEG-USR-030, LEG-USR-033-035, LEG-USR-045-050, LEG-USR-055-056, LEG-USR-058, LEG-USR-068, LEG-USR-079, LEG-USR-083, LEG-USR-088, LEG-USR-091, LEG-USR-096-099.	DL
USR-026	Consumer	C4, C27	TRUSTEE should support data that contain meta-data, text based included.	Desirable Requirement	LEG-TSD-051-062, LEG-TSD-069, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001, LEG-USR-015, LEG-USR-026, LEG-USR-079, LEG-USR-083, LEG-USR-088, LEG-USR-091.	DL, SSI-HE
USR-027	Consumer	C66	TRUSTEE should include good quality metadata	Desirable Requirement	LEG-TSD-026, LEG-TSD-028, LEG-TSD-033, LEG-TSD-069, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001, LEG-USR-015, LEG-USR-026, LEG-USR-079, LEG-USR-083, LEG-USR-088, LEG-USR-091.	DL, SSI-HE
USR-028	Consumer	C22, C25, C62, C72	TRUSTEE should automatically align different data formats coming from different companies and heterogenous systems or repositories	Desirable Requirement	LEG-TSD-013-016, LEG-TSD-018, LEG-TSD-021, LEG-TSD-069-071,	LEG-USR-001, LEG-USR-002, LEG-USR-005-012, LEG-USR-014,	DL, KR

					LEG-TSD-075, LEG-TSD-085, LEG-TSD-095, LEG-TSD-103.	LEG-USR-024, LEG-USR-026-027, LEG-USR-055-058, LEG-USR-060-061, LEG-USR-065-066, LEG-USR-068, LEG-USR-071-073, LEG-USR-079, LEG-USR-083, LEG-USR-088, LEG-USR-091.	
USR-029	Data Provider, Consumer, Model Provider, Developer	DP2, DP3, DP4, DP8, C81, MP5, MP27, MP62, D85	TRUSTEE should offer collaboration across multi-disciplinary data domains	Desirable Requirement	LEG-TSD-012-014, LEG-TSD-036-039, LEG-TSD-051-062, LEG-TSD-064, LEG-TSD-066, LEG-TSD-068-71, LEG-TSD-074-076, LEG-TSD-077, LEG-TSD-079, LEG-TSD-081-083, LEG-TSD-085-086, LEG-TSD-103, LEG-TSD-111-113.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-066, LEG-USR-068-069 LEG-USR-070-072, LEG-USR-079-080, LEG-USR-083, LEG-USR-091, LEG-USR-093-099.	DL, HEDF, KR, DPIA, STM
USR-030	Consumer	C52	TRUSTEE should include multiple querying functions	Desirable Requirement	LEG-TSD-069, LEG-TSD-079, LEG-TSD-095, LEG-TSD-103.	LEG-USR-001, LEG-USR-015, LEG-USR-026, LEG-USR-055-056, LEG-USR-058, LEG-USR-068, LEG-USR-079, LEG-USR-083, LEG-USR-086, LEG-USR-091, LEG-USR-096-099.	KR
USR-031	Consumer	C11	TRUSTEE should explain how the AI model process generates the additional knowledge extracted from data	Desirable Requirement	LEG-TSD-012, LEG-TSD-014, LEG-TSD-017-018, LEG-TSD-033, LEG-TSD-043, LEG-TSD-047, LEG-TSD-049-050,	LEG-USR-001, LEG-USR-014, LEG-USR-018, LEG-USR-024-025, LEG-USR-034-050, LEG-USR-059, LEG-USR-061-062,	AIMaaS (Explainable AI)

					LEG-TSD-071, LEG-TSD-085, LEG-TSD-092-103.	LEG-USR-066, LEG-USR-068-070, LEG-USR-071, LEG-USR-064, LEG-USR-079, LEG-USR-081-093.	
USR-032	Consumer	C6, C77	TRUSTEE should enable query and align datasets with other data from different domains	Desirable Requirement	LEG-TSD-012-014, LEG-TSD-036-039, LEG-TSD-051-062, LEG-TSD-064, LEG-TSD-066, LEG-TSD-068-071, LEG-TSD-074-075, LEG-TSD-077, LEG-TSD-079, LEG-TSD-082, LEG-TSD-085-086, LEG-TSD-103, LEG-TSD-111-113.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-066, LEG-USR-068-069, LEG-USR-070-072, LEG-USR-079-080, LEG-USR-083, LEG-USR-091, LEG-USR-093-099.	HEDF, DL, SSI-HE
USR-033	Consumer	C18, C39	TRUSTEE should enable computations and analysis over the accessed data	Desirable Requirement	LEG-TSD-012-017 LEG-TSD-051-062, LEG-TSD-066, LEG-TSD-068-076, LEG-TSD-080, LEG-TSD-085, LEG-TSD-092-103, LEG-TSD-111-113.	LEG-USR-001-053, LEG-USR-057-058, LEG-USR-059-067, LEG-USR-068-074, LEG-USR-079, LEG-USR-081-093, LEG-USR-096-099.	HEDF, DPIA, STM
USR-034	Data Provider	DP18	The raw data should be maintained in the data provider's database	Desirable Requirement	LEG-TSD-012-019, LEG-TSD-021, LEG-TSD-042-050, LEG-TSD-051-062, LEG-TSD-070, LEG-TSD-074, LEG-TSD-077, LEG-TSD-086, LEG-TSD-109-110.	LEG-USR-001-029, LEG-USR-034-053, LEG-USR-055-058, LEG-USR-059-080, LEG-USR-094-099.	DL
USR-035	Data Provider	DP25	TRUSTEE should support easy transform/reformat of data	Desirable Requirement	LEG-TSD-012-017, LEG-TSD-019, LEG-TSD-055-057, LEG-TSD-068-070,	LEG-USR-001-003, LEG-USR-005, LEG-USR-014, LEG-USR-016,	DL, HEDF

					LEG-TSD-074, LEG-TSD-079, LEG-TSD-082-083, LEG-TSD-085, LEG-TSD-095, LEG-TSD-111.	LEG-USR-018, LEG-USR-024, LEG-USR-026-028, LEG-USR-034-050, LEG-USR-057-058, LEG-USR-059-062, LEG-USR-065-066, LEG-USR-068 LEG-USR-070-073, LEG-USR-079, LEG-USR-083, LEG-USR-086-088, LEG-USR-094, LEG-USR-096-099.	
USR-036	Consumer	C4	TRUSTEE should allow efficient data querying	Desirable Requirement	LEG-TSD-001-037, LEG-TSD-069, LEG-TSD-073-074, LEG-TSD-079, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-095-096, LEG-TSD-098-099, LEG-TSD-103, ~~LEG-TSD-107,~~ LEG-TSD-111-113.	LEG-USR-001-053, LEG-USR-057-058, LEG-USR-059-080, LEG-USR-081-091, LEG-USR-094, LEG-USR-096-099.	HEDF, DL, SSI-HE, DPIA, STM, KR
USR-037	Consumer	C37	TRUSTEE should be able to assess the compliance of an AI model to policies, regulations, and standards	Desirable Requirement	LEG-TSD-001-037, LEG-TSD-092-103.	LEG-USR-001-037, LEG-USR-039-053, LEG-USR-081-093. LEG-USR-098.	TAI-SDF
USR-038	Data Provider	DP3	TRUSTEE should enable transfer of learning without IPR and data privacy issues	Desirable Requirement	LEG-TSD-001-037, LEG-TSD-043-050, LEG-TSD-055-059, LEG-TSD-068-070, LEG-TSD-074, LEG-TSD-077, LEG-TSD-080, LEG-TSD-084, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-092-93, LEG-TSD-095,	LEG-USR-001-053, LEG-USR-054, LEG-USR-055-058, LEG-USR-059-080, LEG-USR-084-085, LEG-USR-086-087, LEG-USR-090, LEG-USR-091-093, LEG-USR-094-099.	DPIA, HEDF, STM

					LEG-TSD-098-099, LEG-TSD-103, LEG-TSD-104, LEG-TSD-106, LEG-TSD-108-113.		
USR-039	Data Provider	DP50, DP62, DP72	TRUSTEE should be secure and free of cybersecurity issues	Desirable Requirement	LEG-TSD-005-007, LEG-TSD-010, LEG-TSD-019, LEG-TSD-021-035, LEG-TSD-042-050, LEG-TSD-051-063, LEG-TSD-072-073, LEG-TSD-075-080, LEG-TSD-087-091, LEG-TSD-092-103.	LEG-USR-001, LEG-USR-024-025, LEG-USR-034-035, LEG-USR-046-049, LEG-USR-051, LEG-USR-058, LEG-USR-068, LEG-USR-071-072, LEG-USR-081-093, LEG-USR-095.	AM, ATR
USR-040	Data Provider	DP75	TRUSTEE should provide protection of personal data	Desirable Requirement	LEG-TSD-001-037, LEG-TSD-042-050, LEG-TSD-054-059, LEG-TSD-092-103, LEG-TSD-109-110.	LEG-USR-001-053, LEG-USR-058, LEG-USR-059-060, LEG-USR-066-072, LEG-USR-074, LEG-USR-079-080, LEG-USR-081-093, LEG-USR-095.	DPIA, STM
USR-041	Consumer	C3	TRUSTEE should allow the process of data without IPR and privacy issues	Desirable Requirement	LEG-TSD-001-037, LEG-TSD-043-050, LEG-TSD-055-059, LEG-TSD-068-070, LEG-TSD-074, LEG-TSD-077, LEG-TSD-080-084, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-092-93, LEG-TSD-095,	LEG-USR-001-053, LEG-USR-054, LEG-USR-055-058, LEG-USR-059-080, LEG-USR-084-085, LEG-USR-086-087, LEG-USR-090, LEG-USR-091-093, LEG-USR-094-099.	HEDF, DPIA, SSI-HE, STM

					LEG-TSD-098-099, LEG-TSD-103, LEG-TSD-104, LEG-TSD-106, LEG-TSD-108-113.		
USR-042	Consumer	C26	TRUSTEE should be compliant with all security requirements and legislations regarding data manipulation and querying	Desirable Requirement	LEG-TSD-001-113.	LEG-USR-001-099.	HEDF, DPIA, SSI-HE, STM
USR-043	Data Provider	DP7	The user should get informed regarding the way that their data are used (accessed, processed, utilized for the training of any model)	Desirable Requirement	LEG-TSD-009, LEG-TSD-013-016, LEG-TSD-018, LEG-TSD-039, LEG-TSD-049, LEG-TSD-054-55, LEG-TSD-058, LEG-TSD-063, LEG-TSD-068-070, LEG-TSD-071, LEG-TSD-076, LEG-TSD-078-080, LEG-TSD-085, LEG-TSD-088-091, LEG-TSD-092-093, LEG-TSD-096, LEG-TSD-098, LEG-TSD-102-103, LEG-TSD-106-109, LEG-TSD-111-113.	LEG-USR-001, LEG-USR-004-005, LEG-USR-007, LEG-USR-009, LEG-USR-011, LEG-USR-014, LEG-USR-016-023, LEG-USR-025-027, LEG-USR-029-050, LEG-USR-053, LEG-USR-057-058, LEG-USR-059-067, LEG-USR-068-078, LEG-USR-079-080, LEG-USR-081-093, LEG-USR-094-099.	STM, ATR, TAI-SDF
USR-044	Consumer	C78	TRUSTEE should allow data that are valid and provided by authenticated users	Desirable Requirement	LEG-TSD-011-013, LEG-TSD-021-022, LEG-TSD-051-062,	LEG-USR-001, LEG-USR-002, LEG-USR-008,	AM

					LEG-TSD-071, LEG-TSD-075, LEG-TSD-085, LEG-TSD-089-090, LEG-TSD-095, LEG-TSD-112-113.	LEG-USR-015, LEG-USR-018-022, LEG-USR-024, LEG-USR-026, LEG-USR-034-035, LEG-USR-045-050, LEG-USR-055-058, LEG-USR-059 LEG-USR-061, LEG-USR-066-067, LEG-USR-068, LEG-USR-072-073, LEG-USR-079-080, LEG-USR-081-084, LEG-USR-086-089, LEG-USR-091, LEG-USR-094, LEG-USR-096-099.	
USR-045	Consumer	C30	TRUSTEE should securely share commercial and private data	Desirable Requirement	LEG-TSD-001-037, LEG-TSD-049, LEG-TSD-055-059, LEG-TSD-066, LEG-TSD-068, LEG-TSD-075-078, LEG-TSD-080, LEG-TSD-084-091, LEG-TSD-095, LEG-TSD-096, LEG-TSD-099, LEG-TSD-103, LEG-TSD-106, LEG-TSD-109, LEG-TSD-111-113.	LEG-USR-001-053, LEG-USR-054, LEG-USR-055-058, LEG-USR-059-080, LEG-USR-084-085, LEG-USR-086-087, LEG-USR-090, LEG-USR-091-093, LEG-USR-094-099.	ATR, DPIA, HEDF, STM
USR-046	Consumer	C74	TRUSTEE should eliminate the correlation of data with personal data	Desirable Requirement	LEG-TSD-004-007, LEG-TSD-013-014, LEG-TSD-019, LEG-TSD-021, LEG-TSD-040, LEG-TSD-055-059, LEG-TSD-070, LEG-TSD-071, LEG-TSD-086, LEG-TSD-095,	LEG-USR-001, LEG-USR-003, LEG-USR-024-025, LEG-USR-028, LEG-USR-058-060, LEG-USR-066-074, LEG-USR-079, LEG-USR-083, LEG-USR-091,	HEDF, DPIA, SSI-HE, STM

					LEG-TSD-103, LEG-TSD-109-110.	LEG-USR-093, LEG-USR-095.	
USR-047	Consumer, Model Provider	C70, MP46, MP77	TRUSTEE should be free, easy-to-use, and offer user-friendly statistical tools	Desirable Requirement	LEG-TSD-001-113.	LEG-USR-001-099.	TRUSTEE Platform, DA (Virtual Assistant), HEDF
USR-048	Data Provider	DP48	TRUSTEE should support the ability to automatically add data and include extra metadata fields to these data if needed	Desirable Requirement	LEG-TSD-012, LEG-TSD-014, LEG-TSD-040, LEG-TSD-056, LEG-TSD-059, LEG-TSD-068-071, LEG-TSD-084, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-112-113.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059, LEG-USR-061, LEG-USR-068, LEG-USR-070-073, LEG-USR-079, LEG-USR-081-084, LEG-USR-086, LEG-USR-091, LEG-USR-094, LEG-USR-096-099.	DA (RPA), DL, SSI-HE, HEDF
USR-049	Consumer	C63	TRUSTEE should perform data querying through user-friendly, simple, and error-free programming languages	Desirable Requirement	LEG-TSD-013-014, LEG-TSD-017-019, LEG-TSD-026-034, LEG-TSD-038-039, LEG-TSD-055-059, LEG-TSD-069-070, LEG-TSD-074-075, LEG-TSD-077, LEG-TSD-085, LEG-TSD-088, LEG-TSD-098, LEG-TSD-102-103, LEG-TSD-106, LEG-TSD-107.	LEG-USR-001, LEG-USR-018, LEG-USR-024, LEG-USR-059, LEG-USR-079, LEG-USR-081-082, LEG-USR-086, LEG-USR-090-091.	DA (RPA)
USR-050	Consumer	C50	TRUSTEE should utilize easy-to-use filters in data querying	Desirable Requirement	LEG-TSD-087-088, LEG-TSD-098, LEG-TSD-102, LEG-TSD-103.	LEG-USR-001, LEG-USR-011, LEG-USR-016, LEG-USR-018, LEG-USR-024, LEG-USR-079, LEG-USR-081-082, LEG-USR-086, LEG-USR-090-091.	DA, HEDF, SSI-HE

USR-051	Consumer	C44, C50	TRUSTEE should allow to perform complex queries without the knowledge of a programming language	Desirable Requirement	LEG-TSD-012-014, LEG-TSD-018, LEG-TSD-022-037, LEG-TSD-042-050, LEG-TSD-087-088, LEG-TSD-095, LEG-TSD-098, LEG-TSD-102, LEG-TSD-103.	LEG-USR-001, LEG-USR-011, LEG-USR-016, LEG-USR-018, LEG-USR-024, LEG-USR-079, LEG-USR-081-082, LEG-USR-086, LEG-USR-090-091.	DA (RPA, Virtual Assistant), OneSS (SDK Toolbox)
USR-052	Consumer	C24	TRUSTEE should extract knowledge without complex analysis methods and tools	Desirable Requirement	LEG-TSD-012-014, LEG-TSD-018, LEG-TSD-022-037, LEG-TSD-042-050, LEG-TSD-087-088, LEG-TSD-095, LEG-TSD-098, LEG-TSD-102, LEG-TSD-103.	LEG-USR-001, LEG-USR-018, LEG-USR-024, LEG-USR-068, LEG-USR-071-074, LEG-USR-079, LEG-USR-081-084, LEG-USR-086, LEG-USR-090-091.	AIMaaS (Core, Explainable AI), DA (RPA), HEDF, KR
USR-053	Consumer, Model Provider	C7, C67, MP75	TRUSTEE should reuse already trained ML and AI models/algorithms	Desirable Requirement	LEG-TSD-001-037, LEG-TSD-092-103, LEG-TSD-111, LEG-TSD-112-113.	LEG-USR-001-053, LEG-USR-081-093, LEG-USR-096-098.	AIMaaS
USR-054	Consumer	C70	TRUSTEE should have easy access and permissions	Desirable Requirement	LEG-TSD-005-007, LEG-TSD-012-018 LEG-TSD-021-037, LEG-TSD-042-050, LEG-TSD-051-062, LEG-TSD-071-073, LEG-TSD-075-076 LEG-TSD-078, LEG-TSD-087-089, LEG-TSD-096, LEG-TSD-111, LEG-TSD-112-113.	LEG-USR-001-003, LEG-USR-007, LEG-USR-009-015, LEG-USR-023-025, LEG-USR-027, LEG-USR-029-033, LEG-034-050, LEG-USR-055-058, LEG-USR-059-062, LEG-USR-064-067, LEG-USR-068-069, LEG-USR-071-074, LEG-USR-076-080, LEG-USR-081-082, LEG-USR-086, LEG-USR-090, LEG-USR-091, LEG-USR-094-099.	DA, AM, STM, DPIA

USR-055	Consumer	C52, C81	TRUSTEE should provide automation techniques	Desirable Requirement	LEG-TSD-001-002, LEG-TSD-007, LEG-TSD-012, LEG-TSD-013-014, LEG-TSD-017-018, LEG-TSD-021-039, LEG-TSD-043-050, LEG-TSD-055, LEG-TSD-057-058, LEG-TSD-063-064, LEG-TSD-066, LEG-TSD-069-073, LEG-TSD-075-078, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-092-103, LEG-TSD-109-111.	LEG-USR-001-002, LEG-USR-005-006, LEG-USR-018-025, LEG-USR-027, LEG-USR-059-062, LEG-USR-064-067, LEG-USR-068-069, LEG-USR-071-074, LEG-USR-076-080, LEG-USR-081-082, LEG-USR-086, LEG-USR-090, LEG-USR-091, LEG-USR-094-099.	DA (RPA)
USR-056	Consumer, Model Provider	C54, MP76	TRUSTEE should offer the ability for automating procedures regarding data consumption and other data processes	Desirable Requirement	LEG-TSD-001-002, LEG-TSD-007, LEG-TSD-012, LEG-TSD-013-014, LEG-TSD-017-018, LEG-TSD-021-039, LEG-TSD-043-050, LEG-TSD-055, LEG-TSD-057-058, LEG-TSD-063-064, LEG-TSD-066, LEG-TSD-069-073, LEG-TSD-075-078, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-092-103, LEG-TSD-109-110.	LEG-USR-001-002, LEG-USR-005-006, LEG-USR-018-025, LEG-USR-027, LEG-USR-034-050, LEG-USR-059-062, LEG-USR-064-067, LEG-USR-068-069, LEG-USR-071-074, LEG-USR-076-080, LEG-USR-081-093, LEG-USR-094-099.	DA (RPA)
USR-057	Data Provider	DP23	TRUSTEE should be an efficient and highly secure platform	Desirable Requirement	LEG-TSD-003 LEG-TSD-005-007, LEG-TSD-010-019, LEG-TSD-021-037, LEG-TSD-042-050, LEG-TSD-051-062, LEG-TSD-071-072, LEG-TSD-076-077,	LEG-USR-001-053, LEG-USR-056, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-081-093, LEG-USR-094-099.	DA, DPIA, STM, TAI-SDF, ATR, AM, SSI-HE, HEDF

					LEG-TSD-079, LEG-TSD-084-085, LEG-TSD-087-091, LEG-TSD-094-103, LEG-TSD-109-110.		
USR-058	Data Provider, Consumer	DP73, C24	TRUSTEE should enable users to access data catalogs, tutorials, and analysis outputs easily	Desirable Requirement	LEG-TSD-001, LEG-TSD-004, LEG-TSD-011, LEG-TSD-014, LEG-TSD-015, LEG-TSD-018, LEG-TSD-020, LEG-TSD-038-039, LEG-TSD-041, LEG-TSD-055, LEG-TSD-058, LEG-TSD-063-079, LEG-TSD-085, LEG-TSD-095, LEG-TSD-111, LEG-TSD-112-113.	LEG-USR-001-053, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-081-093, LEG-USR-094-099.	OneSS (Wiki), DL, KR, DA
USR-059	Consumer, Developer	C49, D86	TRUSTEE should be user-friendly	Desirable Requirement	LEG-TSD-012-016, LEG-TSD-017-018, LEG-TSD-021-022, LEG-TSD-035, LEG-TSD-037-039, LEG-TSD-049, LEG-TSD-055-059, LEG-TSD-071, LEG-TSD-073, LEG-TSD-081-082, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-098, LEG-TSD-102, LEG-TSD-103, LEG-TSD-106, LEG-TSD-107-110, LEG-TSD-111.	LEG-USR-001-053, LEG-USR-055-058, LEG-USR-059-067, LEG-USR-068-080, LEG-USR-081-093, LEG-USR-094-099.	DA, STM, AIMaaS, TAI-SDF, DPIA
USR-060	Consumer	C57	TRUSTEE should show the types of data that are available for access	Desirable Requirement	LEG-TSD-012-016, LEG-TSD-017-018, LEG-TSD-021-022,	LEG-USR-001-003, LEG-USR-007, LEG-USR-014,	DA, DL

					LEG-TSD-035, LEG-TSD-037-039, LEG-TSD-049, LEG-TSD-055-059, LEG-TSD-071, LEG-TSD-073, LEG-TSD-082, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-098, LEG-TSD-102, LEG-TSD-103, LEG-TSD-106, LEG-TSD-107-110, LEG-TSD-111.	LEG-USR-016, LEG-USR-018, LEG-USR-024, LEG-USR-027-028, LEG-USR-055-058, LEG-USR-059, LEG-USR-061, LEG-USR-065-066, LEG-USR-071-072, LEG-USR-074, LEG-USR-079, LEG-USR-083, LEG-USR-094-099.	
USR-061	Model Provider, Consumer, Developer	C49, MP15, MP64, D91	TRUSTEE should have access to support regarding platform functioning	Desirable Requirement	LEG-TSD-007, LEG-TSD-012-017, LEG-TSD-073, LEG-TSD-076-077, LEG-TSD-085, LEG-TSD-088, LEG-TSD-098, LEG-TSD-102, LEG-TSD-103, LEG-TSD-108, LEG-TSD-111.	LEG-USR-001, LEG-USR-012, LEG-USR-018, LEG-USR-024, LEG-USR-025, LEG-USR-034-050, LEG-USR-052, LEG-USR-054, LEG-USR-055-056, LEG-USR-066, LEG-USR-068, LEG-USR-073-074, LEG-USR-079, LEG-USR-081-093, LEG-USR-094-095.	DA (Virtual Assistant), OneSS (Wiki)
USR-062	Consumer	C41	TRUSTEE should have a user-friendly validation and presentation of data analysis results	Desirable Requirement	LEG-TSD-012-016, LEG-TSD-017-018, LEG-TSD-021-022, LEG-TSD-035, LEG-TSD-037-039, LEG-TSD-049, LEG-TSD-055-059, LEG-TSD-071, LEG-TSD-073, LEG-TSD-082, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-098,	LEG-USR-001, LEG-USR-006, LEG-USR-018, LEG-USR-019, LEG-USR-024-025, LEG-USR-055-056, LEG-USR-065-066, LEG-USR-071-073, LEG-USR-079-080, LEG-USR-081-093, LEG-USR-094-095.	DA, AIMaaS (Core, Explainable AI)

					LEG-TSD-102, LEG-TSD-103, LEG-TSD-106, LEG-TSD-107-110, LEG-TSD-111.		
USR-063	Developer	D90	TRUSTEE should have example applications and tutorials for what is offering to be used by end-users	Desirable Requirement	LEG-TSD-007, LEG-TSD-012-017, LEG-TSD-071, LEG-TSD-073, LEG-TSD-076-077, LEG-TSD-081, LEG-TSD-085, LEG-TSD-088, LEG-TSD-098, LEG-TSD-102, LEG-TSD-103, LEG-TSD-108, LEG-TSD-111.	LEG-USR-001, LEG-USR-006, LEG-USR-018, LEG-USR-019, LEG-USR-024-025, LEG-USR-055-056, LEG-USR-065-066, LEG-USR-071-073, LEG-USR-079-080, LEG-USR-081-092, LEG-USR-094-095.	OneSS (Wiki, SDK Toolbox), DA
USR-064	Consumer, Model Provider	C9, C21, C38, C59, C68, C79, MP6, MP7, MP17, MP18, MP28, MP78	TRUSTEE should enable users to provide their AI/ML algorithms for others to use them	Desirable Requirement	LEG-TSD-092-103, LEG-TSD-111, LEG-TSD-112-113.	LEG-USR-019-023, LEG-USR-034-037, LEG-USR-039-041, LEG-USR-081-093, LEG-USR-094-099.	HEDF, ATR, SSI-HE, STM
USR-065	Consumer	C29	TRUSTEE could enable sharing of data in a granular manner	Optional Requirement			
USR-066	Developer	D84, D90	TRUSTEE could utilize extensions for SDK integration in other tools	Optional Requirement			OneSS (SDK Toolbox)
USR-067	Consumer, Model Provider	C43, MP22	TRUSTEE should eliminate the mandatory change of file formats (-universal data file format)	Optional Requirement			DL, HEDF

USR-068	Consumer, Model Provider	C14, MP4	TRUSTEE will not support easy and fast data processing with defined data analysis key performance indicators	Possible Enhancement	Future		AIMaaS, HEDF
USR-069	Consumer	C2	TRUSTEE will not include real-world data	Possible Enhancement	Future		DL
USR-070	Consumer	C2, C29, C49, C66	TRUSTEE will not include high quality data	Possible Enhancement	Future		DL
USR-071	Consumer	C3	TRUSTEE will not include available and reliable data	Possible Enhancement	Future		DL
USR-072	Developer	D86	TRUSTEE will not include free-access data	Possible Enhancement	Future		
USR-073	Consumer	C66	TRUSTEE will not include complete data	Possible Enhancement	Future		DL

USR-074	Model Provider	MP23, MP45	TRUSTEE will not/could check data consistency	Possible Enhancement / Optional Requirement			DL
USR-075	Model Provider	MP3, MP25, MP26	TRUSTEE will not include data quality diagnostic tools	Possible Enhancement			DL, HEDF
USR-076	Data Provider	DP2	TRUSTEE could compress big data in an efficient and fast way	Optional Requirement			
USR-077	Model Provider	MP24	TRUSTEE could support manipulation of data with decimal conversions	Optional Requirement			HEDF
USR-078	Consumer, Model Provider	C69, MP63	TRUSTEE could apply common semantic ontologies to heterogeneous data	Optional Requirement			DL
USR-079	Model Provider	MP72	TRUSTEE could have more data parameter options/fields in a database	Optional Requirement			Data will be stored in the data provider premises
USR-080	Consumer	C52	TRUSTEE will not include automatic identification of the optimal queries	Possible Enhancement			KR
USR-081	Consumer, Model Provider	C2, C14, MP2	TRUSTEE could offer generation of artificial data based on real-world diversity	Optional Requirement			AIMaaS, DL, HEDF

USR-082	Consumer	C19	TRUSTEE could provide information on the way a website/portal collects data from various sources	Optional Requirement			DA (Virtual Assistant), STM
USR-083	Consumer	C76	TRUSTEE will not contain a centralized database	Possible Future Enhancement			DL, KR
USR-084	Consumer	C5	TRUSTEE could enable fast pre-processing of datasets	Optional Requirement			DL, SSI-HE
USR-085	Consumer	C36	TRUSTEE could support improved AI-based analysis methods and tools	Optional Requirement			AIMaaS
USR-086	Model Provider	MP47	TRUSTEE could automatically check consistency of data from different sources and automatically fix the mistakes in these data	Optional Requirement			DL, SSI-HE
USR-087	Consumer	C25	TRUSTEE could allow easy and fast modification of data	Optional Requirement			HEDF, DL
USR-088	Consumer	C12, C14, C80	TRUSTEE will not support big data modification with ease and at runtime	Possible Future Enhancement			HEDF, DL
USR-089	Consumer	C13	TRUSTEE could allow data computations with low computational power	Optional Requirement			AIMaaS
USR-090	Consumer	C61	TRUSTEE could support the utilization of data from healthcare devices	Optional Requirement			DL

USR-091	Consumer	C8	TRUSTEE could assess and monitor the performance of the AI models regarding fairnerss, explainability, robustness, and security	Optional Requirement			TAI-SDF
USR-092	Consumer	C20, C40	TRUSTEE could enable validation of AI support decisions	Optional Requirement			AIMaaS (Explainable AI)
USR-093	Consumer, Model Provider, Developer	C58, C67, C81, D91	TRUSTEE could support fully trustworthy AI/ML models	Optional Requirement			TAI-SDF, AIMaaS (Explainable AI)
USR-094	Data Provider	DP5	TRUSTEE could automatically synchronize datasets while keeping their integrity	Optional Requirement			STM, ATR, DL, HEDF
USR-095	Consumer	C40	TRUSTEE could offer validation of data manipulation results	Optional Requirement			AIMaaS (Explainable AI), HEDF
USR-096	Data Provider, Consumer	DP46, C60	TRUSTEE will not automatically evaluate data entries of users	Possible Future Enhancement			DL, HEDF
USR-097	Consumer	C70	TRUSTEE will not support unbiased basic services operators	Possible Future Enhancement			AIMaaS, TAI-SDF
USR-098	Data Provider	DP29	TRUSTEE will not define functions to validate the entry data easier	Possible Future Enhancement			SSI-HE, HEDF

USR-099	Consumer	C25	TRUSTEE will not support fast access of data	Possible Future Enhancement			DL, SSI-HE, HEDF
USR-100	Data Provider	DP20	The user could be able to decide the level of access they are willing to provide for their data	Optional Requirement			DPIA, STM, AM
USR-101	Data Provider	DP47	TRUSTEE will not automatically anonymize data	Possible Future Enhancement			HEDF
USR-102	Data Provider	DP74	TRUSTEE will not automatically encrypt and secure data	Possible Future Enhancement			DL, HEDF
USR-103	Data Provider	DP22	TRUSTEE will not be confident that cyber security attacks will not result in taking control of big energy resource data	Possible Future Enhancement			AM, STM, DL
USR-104	Data Provider	DP45	TRUSTEE could be a secure platform with waivers from local IRB and privacy policies	Optional Requirement			DL, HEDF, SSI-HE
USR-105	Data Provider	DP26, DP28	TRUSTEE will not overcome business privacy policies	Possible Future Enhancement			TRUSTEE Platform
USR-106	Model Provider	MP44	TRUSTEE will not support easier and at lower cost sustainability of the collected health data	Possible Future Enhancement			DL

USR-107	Data Provider	DP44	Through TRUSTEE data providers could be able to avoid pseudonymization and be able to share data	Optional Requirement			HEDF, SSI-HE, DPIA, STM
USR-108	Data Provider, Consumer, Model Provider, Developer	DP45, C65, MP46, D86	TRUSTEE will not be open or free	Possible Future Enhancement			TRUSTEE Platform (as a whole), OneSS (SDK Toolbox)
USR-109	Developer	D83, D86	TRUSTEE will not contain good inbuilt visualization tools and APIs	Possible Future Enhancement			DA, AIMaaS (Core, Explainable AI), OneSS (SDK Toolbox)
USR-110	Consumer	C17, C25, C28, C64	TRUSTEE could allow easy access of data	Optional Requirement			HEDF, STM
USR-111	Data Provider	DP4	TRUSTEE could annotate data fast and without manual effort	Optional Requirement			HEDF, KR
USR-112	Data Provider	DP27	TRUSTEE will not overcome the issues relating to data collection, sharing business model, and marketing the availability of data	Possible Future Enhancement			DL, HEDF, SSI-HE, DPIA, STM

USR-113	Data Provider	DP27	TRUSTEE could overcome the issues relating to automatic setup of data flows	Optional Requirement			DA (RPA), DL, SSI-HE, HEDF
USR-114	Consumer	C42, C81	TRUSTEE could utilize the same algorithms to manipulate different types of data	Optional Requirement			HEDF, AIMaaS, DL
USR-115	Consumer	C82	TRUSTEE could process valid large volume datasets without the need for specialized infrastructures	Optional Requirement			TRUSTEE Platform, HEDF, AIMaaS, DL
USR-116	Data Provider	DP44	TRUSTEE will not encrypt data easily without having performance overhead or increasing data storage	Possible Future Enhancement			HEDF
USR-117	Model Provider	MP71	TRUSTEE will not be able to personalize the code for training data generation easier	Possible Future Enhancement			TRUSTEE Platform, AIMaaS
USR-118	Model Provider	MP74	TRUSTEE could enable to reach and communicate with more specialists in a specific field	Optional Requirement			DA (Virtual Assistant), KR

ANNEX 5: LEGAL REQUIREMENTS

5.1 LEGAL REQUIREMENTS FOR TRUSTEE END-USERS

Req ID	Personas (Data Provider, Consumer, Model Provider, Developer)	Requirement	Priority (Must-have. Mandatory Requirement, Should-have. Desirable Requirement, Could-have. Optional Requirement, Will-not-have. Possible Future Enhancement)	Architecture Building Block (that could address the requirement)	Legal Act or Relevant Guideline/Standard	Status of the Legal Act or Relevant Guideline/Standard	Comments

LEG-USR-001	Data Provider / Data Consumer / Model Provider	The Data Provider / Consumer / Model Provider providing the data via TRUSTEE needs to consider whether the data are personal or not and indicate it in the STM and in DPIA phase.	Mandatory Requirement	Results to be included in STM and DPIA	GDPR, Article 2 and 4	Adopted, entered into force and applicable	To be considered even before using TRUSTEE.
REQUIREMENTS APPLICABLE TO PERSONAL DATA							
LEG-USR-002	Data Provider / Consumer	The Data Provider / Consumer must indicate appropriate legal basis for processing and sharing the data via TRUSTEE. It is not to be decided by the TRUSTEE system providers.	Mandatory Requirement	STM, DPIA	GDPR Article 5(1)(a)	Adopted, entered into force and applicable	
LEG-USR-003	Data Provider / Consumer	The Data Provider / Consumer should share the data in pseudonymised, or, where possible, in anonymised form, to ensure data minimisation and purpose limitation.	Mandatory Requirement	HEDF	GDPR Article 5(1)(c), Article 32(1), Article 89(1))		To be addressed either before sharing the data on TRUSTEE (anonymisation), or by applying HE (pseudonymisation by hashing the data).
LEG-USR-004	Data Provider / Consumer	Where Data Provider and Consumer process the data for joint purposes by the means provided on TRUSTEE, they should agree on a joint controllership agreement.	Mandatory Requirement	STM	GDPR, Article 26		Mandatory for the users, however, STM could allow for signing such an agreement but it could be also signed outside of TRUSTEE, if required.
LEG-USR-005	Data Provider / Consumer	The Data Provider / Consumer, where considered as necessary, must sign a Data Processing Agreement with TRUSTEE system providers as data processors (when real data are processed).	Mandatory Requirement	STM, ATR	GDPR, Article 28		
LEG-USR-006	Data Provider/ Consumer	Data Provider must provide data subjects with the information that their data will be processed on TRUSTEE and share categories of recipients of data (i.e. categories of Consumers). Consumer should provide the information about processing the data to data subjects, if feasible.	Mandatory Requirement	N/A	GDPR Article 5(1)(a), 12, 13 and 14		Obligation to be realised outside of the scope of TRUSTEE solution.

LEG-USR-007	Data Provider	Data Provider must indicate for what purposes data can be processed by the Consumers. Moreover, the Data Provider and Consumer should include in the data sharing agreement the following information: a) the parties identification and information on the participants of data sharing, b) the purpose of data sharing defined in a clear manner, i.e. description of the aims for which data are shared, why sharing is necessary to achieve those goals and what benefits to individuals and/or to society the parties hope to bring, c) what data are going to be shared d) lawful basis for data sharing (both parties need to clarify it for themselves – Data Providers the basis for making data available, and Consumers for accessing the data) e) information if there is any type of special categories of data that would be shared f) procedures allowing for achieving compliance with the rights of individuals (data subjects) g) governance arrangements between the parties	Mandatory Requirement	STM, DPIA	GDPR, Article 5(1)(b), 6(1), UK ICO data sharing code of practice (see legal framework)	
LEG-USR-008	Data Provider	Data Provider must indicate his/her legal basis for the activity of making data available to Consumers via TRUSTEE.	Mandatory Requirement	STM, DPIA	GDPR, Article 6	
LEG-USR-009	Consumer	Consumer must provide the information concerning processing the data (whenever the Consumer processes personal (encrypted) data provided by the Data Provider) - whether it will correspond with the purposes indicated by Data Provider or not.	Mandatory Requirement	STM, DPIA	GDPR, Article 5(1)(b)	
LEG-USR-010	Data Provider	Data Provider who wants to allow accessing data on TRUSTEE (i.e. will process the data beyond initial purposes) must conduct compatibility test with the initial purpose. In case the purposes are not compatible, Data Provider should abstain from processing the data or provide another lawful basis and inform data subjects about new processing, where possible. The outcome of the compatibility assessment should be indicated by Data Provider in STM and in the context of the DPIA.	Mandatory Requirement	STM, DPIA - both for compatibility test	GDPR, Article 6(4)	Informing data subjects about new processing is binding for Consumer, but to be fulfilled outside of TRUSTEE.

LEG-USR-011	Consumer	Consumer should indicate legal basis for processing of personal, HE encrypted data, provided by Data Provider, and, where applicable, indicate exemption from Article 9 for processing of special categories of data (such as data concerning health, genetic data).	Mandatory Requirement	STM, DPIA	GDPR, Article 6 and Article 9		Obligation to be realised on TRUSTEE or outside of the scope of TRUSTEE solution.
LEG-USR-012	Data Provider	If other purpose is processing for scientific research based on the EU or national law, Data Provider must indicate that. In that case purpose is considered compatible by default, and appropriate technical and organisational safeguards (measures) must be applied.	Mandatory Requirement	STM, DPIA, all other ABBs as technical and organisational safeguards	GDPR, Article 5(1)(c), Article 89(1)		
LEG-USR-013	Data Provider / Consumer	The outcome of the compatibility assessment should be indicated by Data Provider, preferably in a data sharing agreement concluded between Data Provider and Consumer.	Mandatory Requirement	STM, DPIA	GDPR, Article 6(4)		
LEG-USR-014	Data Provider / Consumer	Data Provider and Consumer must process the data only necessary for a specified purpose. Specification should be included in the terms of use of the data (STM) and in the DPIA, and be lawful.	Mandatory Requirement	STM, DPIA	GDPR; Article 5(1)(a), 5(1)(c), 6		
LEG-USR-015	Data Provider	Data Provider must ensure that the data provided onto TRUSTEE are accurate and should indicate it in the contract signed via TRUSTEE.	Mandatory Requirement	STM	GDPR, Article 5(1)(d)		General requirement applicable also prior and outside of TRUSTEE
LEG-USR-016	Data Provider	When a data subject exercises its right of erasure and/or restriction to the processing, Data Provider must ensure to remove the data from TRUSTEE / stop sharing the data of data subjects when requested by a data subject, required by the data protection laws and when there is no legal basis for further sharing and processing of the data.	Mandatory Requirement	ALL ABBs should support this action	GDPR, Article 5(1)(e), 17, 18		
LEG-USR-017	Data Provider / Consumer	Data Provider and Consumer must adhere to the storage limits that they are abide by.	Mandatory Requirement	N/A	GDPR, Article 5(1)(e)		General requirement applicable also prior and outside of TRUSTEE

LEG-USR-018	Data Provider / Consumer	Data Provider and Consumer must answer the data subject's rights: • Information about the data processing (Article 12-14) • Access (Article 15) • Rectification (Article 16) • Erasure (Article 17) • Restriction on processing (Article 18) • Portability (Article 20) • Object (Article 21) • Not to be subject to a decision based solely on automated processing, including profiling (Article 22) • Lodge a complaint with the supervisory authority (Article 13).	Mandatory Requirement	ALL blocks must support this action	GDPR, Article 12-21		Binding even before and outside of TRUSTEE.
LEG-USR-019	Data Provider / Consumer	Data Provider and Consumer must indicate whether application of AI models offered by TRUSTEE will lead to any automated individual decision-making about a data subject whose data are subject to processing.	Mandatory Requirement	DPIA, STM, TAI-SDF	GDPR; Article 22(1)		
LEG-USR-020	Data Provider / Consumer	Data Provider and Consumer must indicate if data subject (whose homomorphically encrypted data is accessed and processed) has been informed about automated individual decision making.	Mandatory Requirement	DPIA, STM, TAI-SDF	GDPR, Article 22(1)		
LEG-USR-021	Data Provider / Consumer	Data Provider and Consumer must provide what is the exception applicable, which allows the automated individual decision making (necessity to perform a contract, authorisation by the EU or national law, explicit consent of a data subject).	Mandatory Requirement	DPIA, STM, TAI-SDF	GDPR, Article 22(2)		
LEG-USR-022	Data Provider / Consumer	For special categories of data, Data Provider and Consumer must indicate whether automated individual decision making is allowed due to an explicit consent of a data subject or due to a substantial public interest, based on the EU or national law.	Mandatory Requirement	DPIA, STM, TAI-SDF	GDPR, Article 22(3)		
LEG-USR-023	Data Provider / Consumer	Data Provider and Consumer applying AI models must perform a DPIA.	Mandatory Requirement	DPIA, TAI-SDF	GDPR; Article 35(1) and WP29 guidelines		

LEG-USR-024	Data Provider / Data consumer	Data Provider and Consumer must use the technical and organisational measures to protect personal data. Data provider and Consumer should use measures offered by TRUSTEE (Homomorphic encryption, SSI, authentication, federated learning, etc.).	Mandatory Requirement	ALL	GDPR, Article 5(1)(f), Article 32		
LEG-USR-025	Data Provider / Consumer	Data Provider and Consumer have to carry out an objective quantitative and qualitative risk assessment of the processing of the data they foresee to conduct. The more severe a potential risk, the higher level of security to mitigate the risk should be.	Mandatory Requirement	DPIA	GDPR, Article 32		
LEG-USR-026	Data Provider Consumer	Data Provider and Consumer should make sure that the data are authentic and accurate, and have not been maliciously or accidentally altered at any stage (when processed, stored or transmitted).	Mandatory Requirement	N/A	GDPR, Article 32		Data quality is to be addressed, but outside of TRUSTEE (prior and during the exchange of the data)
LEG-USR-027	All	All actors should be able to access only the data they actually need.	Mandatory Requirement	AM, HEDF, STM, DPIA	GDPR, Article 5, 32(1), CNIL GDPR developer guide		
LEG-USR-028	Data Provider / Consumer	Data Provider and Consumer should always take into account the risk of re-identification of a data subject and should not attempt to re-identify a data subject when it is not necessary or forbidden by the Data Provider. Appropriate statements referring to that should be included in the agreements.	Mandatory Requirement	STM, DPIA	GDPR, Article 5(1)(c)		Applicable not only on, but also outside of TRUSTEE.
LEG-USR-029	Consumer	Consumer should adhere to the confidentiality standards required by Data Provider.	Mandatory Requirement	STM, HEDF	GDPR, Article 5(1)(f), Article 32		
LEG-USR-030	Data Provider	Data Provider may indicate to what countries the data can be made available.	Mandatory Requirement	STM, ALL ABBs should support such limitation	GDPR, Article 45-49 GDPR, CNIL GDPR developer guide		
LEG-USR-031	Data Provider	Data Provider may indicate that the data can be provided only to the Consumers from the countries with an adequate level of protection, as deemed by the European Commission.	Mandatory Requirement	STM, ALL ABBs should support execution of such limitation	GDPR, Article 45		

LEG-USR-032	Data Provider / Consumer	Data Provider should assess the legislation of the third country where a Consumer is based and they would enable access to the data to. Data Provider may require that the Consumer from non-EU country provides sufficient safeguards to protect the data in and after transfer. This should be included in the agreement and assessment report. Those safeguards may be: a. a legally binding and enforceable instrument between public authorities or bodies; b. binding corporate rules in accordance with Article 47 GDPR; c. standard contract clauses adopted by the Commission; d. standard contract clauses adopted by a supervisory authority and approved by the Commission; e. an approved code of conduct together with binding and enforceable commitments to apply safeguards; or f. an approved certification mechanism together with binding and enforceable commitments to apply safeguards.	Mandatory Requirement	STM, DPIA	GDPR, Article 46, EDPB, Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data (2021)	
LEG-USR-033	Data Provider / Consumer	Data Provider may require the Consumer located in non-EU country to sign standard contractual clauses / data sharing agreement before obtaining the access to the data.	Mandatory Requirement	STM	GDPR, Article 46	STM could support such solution
LEG-USR-034	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must go through a Data Protection Impact Assessment (DPIA) prior to processing the data if processing is likely resulting in a high risk to the rights and freedoms of the people whose data they process, as well as if two of req 36-44 materialise.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA	
LEG-USR-035	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must identify risks to data processing and their mitigation and indicate it in the DPIA before granting actual access to (homomorphically encrypted) data.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA	

LEG-USR-036	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if the envisioned processing involves evaluation or scoring, profiling and predicting activities.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-037	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if they envisage automated individual decision making with legal or similar significant effects to a person.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-038	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if they envisage systematic monitoring.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-039	Data Provider Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if they process special categories of data, such as, e.g. health data or genetic data.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-040	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if they envisage processing of the data on a large scale.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-041	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if they envisage matching or combining datasets.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-042	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if they envisage processing of the data of vulnerable individuals - e.g. patients or minors.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-043	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if they envisage data transfer across borders outside the European Union.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		

LEG-USR-044	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a DPIA if the processing in itself “prevents data subjects from exercising a right or using a service or a contract”.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-045	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) should conduct a DPIA in all the cases where it is not clear whether a DPIA is required.	Mandatory Requirement	DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
LEG-USR-046	Data Provider / Consumer	In the DPIA, Data Provider / Consumer who will process personal data (also HE encrypted) must provide a systematic description of the envisaged processing operations and the purpose of processing.	Mandatory Requirement	DPIA	GDPR, Article 35 (7); WP29 guidelines on a DPIA		
LEG-USR-047	Data Provider / Consumer	In the DPIA, Data Provider / Consumer who will process personal data (also HE encrypted) must conduct a proportionality and necessity assessment of processing operations in relation to the purposes.	Mandatory Requirement	DPIA	GDPR, Article 35 (7); WP29 guidelines on a DPIA		
LEG-USR-048	Data Provider / Consumer	In the DPIA, Data Provider / Consumer who will process personal data (also HE encrypted) must conduct an assessment of the risks to the rights and freedoms of individuals whose data they process.	Mandatory Requirement	DPIA	GDPR, Article 35 (7); WP29 guidelines on a DPIA		
LEG-USR-049	Data Provider / Consumer	In the DPIA, Data Provider / Consumer who will process personal data (also He encrypted) must indicate the measures to address the risks, such as appropriate safeguards and security measures.	Mandatory Requirement	DPIA	GDPR, Article 35 (7); WP29 guidelines on a DPIA		
LEG-USR-050	Data Provider / Consumer	Data Provider / Consumer who will process personal data (also HE encrypted) shall take into account applicable national laws and lists of the activities which require a Data Protection Impact Assessment.	Mandatory Requirement	DPIA	GDPR, Article 35		
LEG-USR-051	Data provider / Consumer	Data Provider / Consumer must inform a supervisory authority if processing, even with security measures, will result in a high risk and/or if they do not apply any measures to mitigate the risk.	Mandatory Requirement	N/A	GDPR, Article 36		Applicable to Data Provider and Consumer, but outside of the TRUSTEE scope.

LEG-USR-052	Data provider / Consumer	Data Provider and Consumer must keep the record of the DPIA.	Mandatory Requirement	STM, DPIA, ATR	GDPR, Article 58(1)(a)		
LEG-USR-053	Data Provider	Data Provider must inform Consumer about any applicable internal policies that should be applicable to data sharing and processing.	Mandatory Requirement	STM, DPIA	GDPR, Article 35, WP29 guidelines on a DPIA		
REQUIREMENTS APPLICABLE TO PERSONAL AND NON-PERSONAL DATA							
LEG-USR-054	Data Provider / Consumer / Model Provider	Data Provider / Consumer / Model Provider should be able to easily port the data when switching to other data service providers, i.e. to other platforms similar to TRUSTEE.	Mandatory Requirement	ALL	Regulation on free flow of non-personal data, Article 3(8)	Adopted, entered into force and applicable	
LEG-USR-055	Data Provider / Consumer / Model Provider	If Data Provider / Consumer / Model Provider is a public body, they should make documents and research data available based on a request, where possible and appropriate, by electronic means, in formats that are open, machine-readable, findable and re-usable, together with their metadata. It does not include all documents listed in ODD, Article 1(2)! Personal data should still be protected according to the GDPR.	Desirable Requirement	STM, ATR, KR,	Open Data Directive, Article 5(1) in conjunction with Article 1(1) and 1(2)		General requirement binding whether the actors are users of TRUSTEE or not. TRUSTEE could facilitate it.
LEG-USR-056	Data Provider / Consumer / Model Provider	Data Provider / Consumer / Model Provider being a public body should make dynamic data available through appropriate technical means immediately after their collection.	Desirable Requirement	STM, ATR, HEDF	Open Data Directive, Article 5(5) in conjunction with Article 2(8)		General requirement binding whether the actors are users of TRUSTEE or not. TRUSTEE could facilitate it.
LEG-USR-057	Data Provider / Model Provider	Data Provider / Model Provider being a public body should not provide access to/make the re-use of documents subject to conditions, unless they are “objective, proportionate, non-discriminatory and justified on grounds of a public interest objective” as well as not being unnecessarily restrictive of the possibilities for re-use. Data Providers / Model Providers may apply a licence that the data consumer must adhere to.	Desirable Requirement	STM	Open Data Directive, Article 8		General requirement binding whether the actors are users of TRUSTEE or not. TRUSTEE could facilitate it.

LEG-USR-058	Data Provider / Model Provider	Data Provider / Model Provider being a public body should make their research data as open as possible, as close as necessary, taking into account any concerns regarding, inter alia, IP rights, confidentiality, security and the protection of personal data. Research data shall be re-usable for commercial or non-commercial purposes in accordance with ODD, insofar as they are publicly funded and researchers, research performing organisations or research funding organisations have already made them publicly available through an institutional or subject-based repository.	Desirable Requirement	HEDF, STM, DPIA	Open Data Directive, Article 10(1)and 10(2)		General requirement binding whether the actors are users of TRUSTEE or not. TRUSTEE could facilitate it.
LEG-USR-059	Data Provider / Model Provider	The use of the data may be made by Data Provider / Model Provider being public sector body conditional on the adherence by the Consumer to a confidentiality obligation prohibiting disclosure of any information jeopardizing the rights and interests of third parties that the data user may have acquired despite the safeguards put in place. This is a general requirement binding public sector bodies which are competent to grant or refuse access for the re-use of data. Thus, TRUSTEE can provide the tools to help facilitate this, if possible.	Desirable Requirement	STM	DGA, Recital 15, Article 5(5)	Law in force, but applicable from 24 September 2023	
LEG-USR-060	Consumer	Consumer re-using the data from a public body is prohibited re-identifying any data subject to whom the data relates and must take technical and operational measures to prevent re-identification and to notify any data breach resulting in the re-identification of the data subjects concerned to the public sector body.	Mandatory Requirement	STM, DPIA, SSI-HE, HEDF, could be addressed by the ToU of TRUSTEE	DGA, Article 5(5)		
LEG-USR-061	Data Provider / Model Provider	Data provider / Model provider may set list of requirements related to intellectual property rights, trade secrets and contractual guarantees to avoid any unlawful or unauthorized access to the non-personal data, and Consumer must adhere to them when processing the data / making use of the model.	Mandatory Requirement	STM	DGA, Recital 20, Article 5(3)(a)(ii), 5(5),(6),(7),(8)		

LEG-USR-062	Consumer	In the case of unauthorised re-use of non-personal data, Consumer must inform legal persons whose rights and interests may be affected.	Mandatory Requirement	STM - could be addressed by the ToU of TRUSTEE, to be accepted by the Consumer	DGA, Article 5(5)		General requirement for Data Consumer, binding also outside of TRUSTEE architecture.
LEG-USR-063	Data Provider	Data Provider being a public sector body must provide assistance to potential re-users/data users in seeking consent of the data subjects or permission from the data holders, where it is feasible without a disproportionate burden on the public sector body. Furthermore, confidential information (e.g. trade secrets) can be disclosed for re-use only with such consent or permission.	Mandatory Requirement	STM - could be addressed by the ToU of TRUSTEE, to be accepted by the Consumer	DGA, Article 5(6)		General requirement for Data Provider, binding also outside of TRUSTEE architecture.
LEG-USR-064	Data Provider	Agreements or other practices pertaining to the re-use of data held by public sector bodies containing selected categories of data which grant exclusive rights or which have as their objective or effect to grant such exclusive rights or to restrict the availability of data for re-use by entities other than the parties to such agreements or other practices shall be prohibited (with exceptions concerning data from Article 3(1) DGA).	Mandatory Requirement	STM - to be considered whether this limitation can be addressed in the agreement manager	DGA, Article 4(1) in conjunction with Article 3(1)		General requirement binding whether the actors are users of TRUSTEE or not. This can also be a term under TRUSTEE Terms of Use where the Data Provider promises not to make such exclusive arrangements while using TRUSTEE and providing access to its data via TRUSTEE, provided that they are public sector bodies.
LEG-USR-065	Data Provider / Model Provider	Data Provider / Model Provider competent to grant or refuse access for the re-use of the data are required to make the conditions for allowing such re-use publicly available. The conditions for re-use shall be non-discriminatory, transparent, proportionate and objectively justified with regard to the categories of data and the purposes of re-use and the nature of the data for which re-use is allowed.	Mandatory Requirement	DA (RPA), STM, to be also addressed by TRUSTEE ToU	DGA, Article 4(2),(3) and (4) DGA, Article 5(1),(2)		
LEG-USR-066	Data Provider / Model Provider	Data Provider / Model Provider being a public sector body must, in accordance with law, ensure that the protected nature of data is preserved.	Mandatory Requirement	STM, HEDF, ATR, DPIA, to be also considered in TRUSTEE ToU	DGA, Article 5(3)		

LEG-USR-067	Data Provider / Model Provider	Data Provider / Model Provider being public sector body may charge fees for allowing the re-use of data from Article 3(1), but these should be transparent, non-discriminatory, proportionate and objectively justified and shall not restrict competition, and shall be limited to the necessary costs related to, inter alia, reproduction, provision, dissemination of data, clearance of rights, anonymization or other form of preparation of personal and commercially confidential data, assistance in seeking consent from data subjects and permission from data holders.	Future enhancement	N/A	DGA, Article 6		Outside of the scope of TRUSTEE
LEG-USR-068	Data Provider / Model Provider	If the Data Provider / Model Provider is going to be a "data holder" in the meaning of the Data Act (see legal framework), they should fulfil the obligations listed in Article 4 Data Act concerning, inter alia, way of making the data available to the user, the form in which the data should be made accessible, contractual agreement which may be concluded between user of a connected device/related service and Data provider, liability clauses, enabling identification through eID wallets, collecting minimum data about the user, protection of trade secrets (including construction of model contractual terms, confidential agreements, strict access protocols, technical standards), sharing of personal data and non-personal data does not apply to information derived or inferred from the data by complex algorithms combining output of multiple sensors.	Desirable Requirement	ALL	Data Act, Article 4 (1)-4(6b)	At the proposal stage, however, relevant for TRUSTEE if adopted. To be monitored for further changes.	
LEG-USR-069	Data Provider	Data Provider may reject a request for data access if it is prohibited by law.	Desirable Requirement	STM	Data Act, Article 3(1a) and 4 (1a)		

LEG-USR-070	Data Provider	Data Provider (if considered a "data holder") must respect the rights of the users of connected devices and related services to share the data with other parties. Data Provider must process the data of users of connected devices (being data subjects) for the purposes specified by the data subject, and in the case when the data subject is not the user and requests access to data, any personal data obtained, collected, or generated by their use of a connected product or related service, and data derived and inferred from that use, shall only be made available by the Data Provider to the third party where there is a valid legal basis under Article 6 (and 9 where relevant) GDPR, with respect to the data subjects rights.	Desirable Requirement	DPIA, STM,	Data Act, Article 5(1)-5(9)		General requirement for Data Provider, binding also outside of TRUSTEE architecture. To be considered if out of the scope of TRUSTEE.
LEG-USR-071	Consumer	<u>If a Consumer is "a data recipient receiving data at the request of the user" in the meaning of the Data Act</u> , has to respect the obligations including (but not limited to): respecting the purpose limitation, legal basis, storage limitation and conditions of data processing agreed with the user, not making the exercise of users' rights difficult, not using the data for profiling of natural persons (unless allowed by the GDPR), not making data available to third parties without informing user and seeking explicit, contractual permission from the user, not making data available to the providers of services who have been designated as gatekeepers (according to the DMA), not using or sharing the data to develop a competing product, not using data to assess economic situation, assets and production methods or use by the data holder (Data Provider), not using the data adversely impacting the security of products, not disregarding trade secrets aspects, not using the data to disrupt sensitive critical infrastructure.	Desirable Requirement	DPIA, STM, HEDEF, AM,	Data Act, Article 6(1)-6(2a)		Binding for Consumer regardless whether it uses or does not use TRUSTEE. TRUSTEE can support fulfilment of the obligations.

LEG-USR-072	Data Provider / Consumer	If Data Provider is considered "data holder" and Consumer "data recipient", and in case the Data provider is required to make the data available to Consumer under Union or national law, they should agree on the modalities for making the data available and base it on fair, reasonable, transparent and non-discriminatory terms. They cannot be required to provide any information beyond what is necessary to verify compliance with the contractual terms agreed for making data available. They must take all necessary legal, organisational and technical measures to ensure the security and integrity of the data transfers. They should also not be obliged to disclose the data being trade secrets (unless required by law).	Desirable Requirement	STM, ATR	Data Act, Article 8 (1), 8(5), 8(5a), 8(6)		
LEG-USR-073	Data Provider	Data Provider (being a "data holder") may apply technical protection measures, <u>including smart contracts and encryption</u> , to prevent unauthorised disclosure of and access to the data, including metadata, and to ensure compliance with Data Act obligations, as well as with the agreed contractual terms for making data available. This protection should not discriminate data recipients (i.e. Consumers).	Desirable Requirement	STM, DPIA, ATR, TAI-SDK, HEF, AM, HEDF	Data Act, Article 11(1)		
LEG-USR-074	Consumer / Data Provider	If a Consumer (being "data recipient" in the meaning of the Data Act) has provided false information or deployed false information to the Data Provider ("data holder"), deployed deceptive or coercive means or abused evident gaps in the technical infrastructure of the Data Provider (data holder) designed to protect the data, has used the data made available for unauthorised purposes, including the development of a competing product or has unlawfully disclosed data to another party, the Consumer (data recipient) shall be liable for the damages to the party suffering from the misuse or disclosure of such data and shall comply without undue delay with the requests of the Data Provider or the trade secret holder when they are not the same legal person to: (a) erase the data made available and any copies thereof;	Desirable Requirement	STM, DPIA, ATR, HEDF, KR, AImaaS, DL, HEDF	Data Act, Article 11(2)		

		(b) end the production, offering, placing on the market or use of goods, derivative data or services produced on the basis of knowledge obtained through such data, or the importation, export or storage of infringing goods for those purposes, and destroy any infringing goods; (c) inform the user (data subject) of the unauthorised use or disclosure of the data and measures taken to put an end to the unauthorised use or disclosure of the data; (c) notify the Data Provider about the disclosure of such data.					
LEG-USR-075	Data Provider / Consumer	Both Data Provider and Consumer cannot unilaterally impose unfair contractual terms concerning the access to and use of data or the liability and remedies for the breach or the termination of data related obligations. Unfair term is defined in Data Act, Article 13(2),(3) and (4).	Desirable Requirement	STM	Data Act, Article 13(1)		General requirement binding whether the actors are users of TRUSTEE or not. STM could facilitate the answer to it.
LEG-USR-076	Data Provider / Consumer	Upon a specified duly justified request limited in time and scope, a Data Provider or a Consumer (if considered a "data holder" in the meaning of the Data Act), that is a legal person shall make non-personal data which are available at the time of the request, including metadata, available to a public sector body or to a Union institution, agency or body demonstrating an exceptional need to use the data requested (defined in Article 15), except for SMEs. Enterprises that fall within the scope of this obligation shall inform their users of the possibility that data may be shared in the case of exceptional circumstances.	N/A	N/A	Data Act, Article 14(1), Article 16(2a)		An obligation which Data Provider and/or Consumer has to address, but addressing it is outside of TRUSTEE.
LEG-USR-077	Data Provider / Consumer	Data Provider / Consumer being data holders should respond to the request from public bodies to access the data according to Article 18 Data Act. Remuneration for access should be not required, unless conditions from Article 20(2) Data Act materialise.	N/A	N/A	Data Act, Article 18		An obligation which Data Provider and/or Consumer has to address, but addressing it is outside of TRUSTEE. Under evaluation if to keep it or not.
LEG-USR-078	Data Provider / Consumer	Consumer being a public sector body or a Union institution, agency or body shall be entitled to	N/A	N/A	Data Act, Article 21(1) and 21(4)		An obligation which Data Provider and/or Consumer has to address,

		share data received with individuals or organisations in view of carrying out scientific research or analytics necessary to fulfil the purpose for which the data was requested, or to national statistical institutes, the members of the European System of Central Banks and Eurostat for the compilation of official statistics. They should notify a Data provider.					but addressing it is outside of TRUSTEE. Under evaluation if to keep it or not.
LEG-USR-079	Data Provider / Model Provider / Developer	Data Provider / Model Provider / Developer being a "participant of data space offering data or data services to other participants" shall comply with the essential requirements to facilitate interoperability of data, data sharing mechanisms and services, defined in Article 28 Data Act: (a) the dataset content, use restrictions, licences, data collection methodology, data quality and uncertainty shall be sufficiently described in a machine-readable format to allow the recipient (Consumer) to find, access and use the data; (b) the data structures, data formats, vocabularies, classification schemes, taxonomies and code lists shall be described in a publicly available and consistent manner; (c) the technical means to access the data, such as application programming interfaces, and their terms of use and quality of service shall be sufficiently described to enable automatic access and transmission of data between parties, including continuously or in real-time in a machine-readable format where that is technically feasible and does not hamper the good functioning of the product; (d) the means to enable the interoperability of contracts for data sharing within their services and activities shall be provided.	Desirable Requirement	STM, DPIA, KR, DA, HEDF	Data Act, Article 28(1)		
LEG-USR-080	Data Provider / Model Provider / Developer	Data Provider / Model Provider / Developer offering smart contract to Consumer in the context of an agreement to make data available shall comply with the essential requirements defined in Article 30 Data Act, concerning robustness and access control, safe termination and interruption, equivalence, protection of confidentiality of trade secrets, but also auditability.	Desirable Requirement	STM, ATR, TAI-SDF	Data Act, Article 30 (EP and EC versions read together)		

LEG-USR-081	Model Provider / Developer	Model Providers and Developers must determine if the AI provided by them is a high risk AI.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 6(1) and (2), Article 6(3) and Annex III	At the proposal stage, however, relevant for TRUSTEE when adopted. To be monitored for further changes.	
LEG-USR-082	Model Provider / Developer	If AI provided by the Model Provider / Developer is high risk, a risk management system must be established by them.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 9		
LEG-USR-083	Model Provider / Developer	If the system provided by Model Provider / Developer is high risk AI, the training, validation and testing data sets must meet a certain quality criteria from Article 10, if the high-risk AI system uses techniques including the training of models with data. The data governance practices must concern in particular: relevant design choices, transparency as regards the original purpose of data collection, data collection process, data preparation operations (annotation, labelling, cleaning, updating, enrichment and aggregation), the formulation of assumptions, an assessment of the availability, quantity and suitability of the needed datasets, examination in view of possible biases, appropriate measures to detect, prevent and mitigate possible biases, identification of data gaps or shortcomings and how they can be addressed. At the level of individual datasets and their combination, the training, validation and testing datasets (incl. labels) shall be relevant, representative, appropriately vetted for errors and complete, have appropriate statistical properties, including, as regards the persons or groups of persons in relation to whom the high-risk AI system is intended to be used.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 10(2)-(5)		
LEG-USR-084	Model Provider / Developer	If AI provided by the Model Provider / Developer is high risk, the events in the AI system must be automatically recorded (logs must be kept), to enable tracing of AI system's functionality and to monitor the situations posing risks.	Desirable Requirement	AIMaaS, TAI-SDF, ATR	AIA, Article 12, Article 20 and Annex IV		

LEG-USR-085	Model Provider / Developer	If AI provided by the Model Provider /Developer is high risk, a technical documentation of the high-risk AI system must be provided before it is placed on the market or put into service, and regularly updated.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 11		
LEG-USR-086	Model Provider / Developer	If AI provided by the Model Provider / Developer is high risk, Model Provider / Developer must ensure that their (AI systems') operation is sufficiently transparent to enable users (Consumers) to interpret the system's output and use it appropriately, by providing instructions in an appropriate digital format and information about the purpose, capabilities and limitations of the AI system in a clear manner. Specific parts of the information are included in Article 13 AIA. Additionally, the AI systems must be designed and developed in a way that they can be effectively overseen by natural persons during the time the AI system is in use. This refers also to appropriate human-machine interface tools.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 13, Article 14, Article 52		
LEG-USR-087	Model Provider / Developer	If AI provided by the Model Provider / Developer is high risk, Model Provider / Developer shall ensure appropriate level of accuracy, robustness, cybersecurity. The levels of accuracy and accuracy metrics shall be declared in the instructions of use. The AI systems must be adaptable against errors, faults and inconsistencies that may occur within the system or the environment in which it operates and withstand any attempts by unauthorised parties to alter the use or performance of the AI system.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 15		
LEG-USR-088	Model Provider / Developer	If AI provided by the Model Provider / Developer is high risk, Model Provider / Developer must establish and document a post-monitoring plan for high-risk AI systems to actively and systematically collect, document and analyse data provided by users or other data collected on the performance of the AI system throughout their lifecycle.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 61		

LEG-USR-089	Model Provider / Developer	Model Provider / Developer must provide appropriate quality management system and implement a conformity assessment procedure for the AI system before it is placed on the market.	Desirable Requirement	AIMaaS, TAI-SDF	AIA, Article 17, Article 19		
LEG-USR-090	Model Provider / Developer	Model Provider / Developer who provides AI system (not only high-risk AI system) must communicate clearly to a user being a natural person that this person interacts with an AI system, unless evident from the context.	Desirable Requirement	AIMaaS, TAI-SDF, DPIA, DA	AIA, Article 52		
LEG-USR-091	Model Provider / Developer Consumer	Model Provider / Developer and Consumer using models must adhere to the <u>requirements for Trustworthy AI: Human Agency and Oversight, Technical Robustness and Safety, Privacy and Data Governance, Transparency, Diversity, Non-discrimination and Fairness, Environmental and Societal Well-Being, Accountability, further described in the legal framework and in socio-ethical requirements section.</u>	Mandatory Requirement	AIMaaS, TAI-SDF, DPIA	EC Ethics guidelines on trustworthy AI, Assessment List TAI (both introduced in legal framework and further detailed in the table on socio-ethical requirements)	Adopted, entered into force and applicable	
LEG-USR-092	Model Provider/ Developer / Consumer	Model Provider/ Developer / Consumer using the AI model should consider conducting fundamental rights impact assessment (in addition to a DPIA) as part of AI impact assessment.	Desirable Requirement	TAI-SDF	See legal framework		

LEG-USR-093	Model Provider / Developer / Consumer	Model Provider and Developer, when developing, and Consumer, when using algorithms for the use in HEALTH domain, must consider, in addition to the requirements above, fulfilling the following: -preserving autonomy of a data subject, - ensure safety, accuracy and efficiency, quality control and improvement (impacting human well-being, human safety and the public interest); - ensure that the system is intelligible and understandable to developers, users and regulators (transparency, explainability and intelligibility); - consider responsibility for the AI system and ensure that AI relies on certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment (responsibility and accountability); - consider different languages, ability and forms of communications, to make AI system as inclusive as possible and non-discriminatory, elimination of unintended bias in the design, implementation and use of the algorithms (inclusiveness and equity); - observe how developed AI answers to communicated expectations and requirements in the context in which it is used.	Mandatory Requirement	TAI-SDF, DPIA, AIMaaS	Ethics and governance of artificial intelligence for health: WHO guidance	
LEG-USR-094	Data Provider / Consumer / Model Provider / Developer	Data sharing agreements concluded via smart contracts on TRUSTEE must be transparent, respect commercial interests of the parties, ensure undistorted competition, minimise data lock-in. The contracts should include (the catalogue is not closed) provisions concerning the scope (nature of data), commercial/business related provisions, provisions concerning control, ownership, access and (re-)usage rights, provisions related to the service providing the data, provisions on limitations stemming from different obligations of Data Providers/ Model Providers/ Developers and Consumers having impact on data sharing and further processing, technical measures for data access and sharing, as well as security measures, liability clauses and provisions on Data Providers' and Consumers' rights such as intellectual property	Mandatory Requirement	STM	Commission Staff Working Document, Guidance on sharing private sector data in the European data economy	

		rights, trade secrets, sensitive commercial information.					
LEG-USR-095	Data Provider / Model Provider / Consumer / Developer	In the context of blockchain usage, Data Provider / Model Provider / Consumer / Developer should make sure that these technologies allow them, as the controllers, to adhere to the data protection and security standards, especially in the context of subcontracting, adherence to the GDPR rules concerning: international transfers of personal data, privacy by design and by default, DPIA, storage limitation, data minimisation and purpose limitation, Technical and organisational measures, recognition and appropriate assignment of the roles in the data processing.	Mandatory Requirement	STM, HEDF, ATR, DPIA	GDPR, NIS Directive, Data Act, contractual law applicable to data provider and data consumer		
LEG-USR-096	Data Provider / Model Provider	Data Provider / Model Provider must indicate the limitations for data processing stemming from their intellectual property rights and sui generis protection of databases (contents of databases), computer programmes, software, and AI/ML models (where this qualifies for any IPR protection). This can be implemented by way of licenses arranged between Data/Model Providers and Consumers.	Mandatory Requirement	STM	Infosoc Directive, Software Directive, Database Directive		

LEG-USR-097	Data Provider / Model Provider	Data Provider / Model Provider must indicate the limitations and obligations binding for Consumers, stemming from the protection from trade secrets law. Any contractual limitations to the further use or disclosure of the trade secret shared as data through TRUSTEE, will continue to apply and any breach of these will be considered as unlawful use or disclosure of the trade secret if the use or disclosure of it is in “breach of a confidentiality agreement or any other duty not to disclose the trade secret” or is in breach of a contractual or any other duty to limit the use of the trade secret.”	Mandatory Requirement	STM	Trade Secret Directive, Article 4(3)		
LEG-USR-098	Data Provider / Model Provider	Data Provider / Model Provider must give truthful information with respect to intellectual property rights held by them on any type of data (where this qualifies for copyright protection), database, software and/or AI/ML models (where this qualifies for any IPR protection) and make necessary attribution to rightholders, where relevant. In case the data provider / model provider is not the rightholder, they must ensure that necessary permissions are obtained before making available any data / model via TRUSTEE.	Mandatory Requirement	STM - could be addressed also in the TRUSTEE ToU	Intellectual property law framework		
LEG-USR-099	Data provider / Consumer	Where applicable, Data Provider should consider if national laws or ethical standards require additional consent and permissions or a specific legal basis for sharing the data with Consumer.	Mandatory Requirement	STM	GDPR, Article 6(3), Article 9(4) Ethical standards in HEALTH (see legal framework)	Adopted, entered into force and applicable	

5.2 LEGAL REQUIREMENTS FOR TRUSTEE SYSTEM DEVELOPERS

Req ID	Requirement	Priority (M: Must-have. Mandatory Requirement, S: Should-have. Desirable Requirement, C: Could-have. Optional Requirement, W: Will-not-have. Possible Future Enhancement)	Architecture Building Block (that could address the requirement)	Legal Act or Relevant Guideline/Standard	Status of the Legal Act or Relevant Guideline /Standard
REQUIREMENTS APPLICABLE TO PERSONAL DATA					
LEG-TSD-001	Should the TRUSTEE system developers and parties maintaining the system at any point handle or process any type of personal data, they should balance their legitimate interest with the fundamental rights and freedoms of the data subjects whose data they will be processing.	Mandatory Requirement	ALL ABBs	Charter of Fundamental Rights of the EU, Art 7 and 8; Article 12 of the Universal Declaration of Human Rights	Law and / or guideline in force and applicable
LEG-TSD-002	TRUSTEE must have the fundamental rights principles at its core, by ensuring that the privacy of the individuals is respected and protected.	Mandatory Requirement	All ABBs	Charter of Fundamental Rights of the EU, Art 7 and 8; Article 12 of the Universal Declaration of Human Rights	
LEG-TSD-003	TRUSTEE must be developed as the system ensuring appropriate and secure processing of personal data. This is relevant especially for: a) the data considered as “special category data”, such as data concerning health or genetic data, as well as b) personal data concerning minors. These data will be processed by the future users of TRUSTEE (Data providers, Data consumers), especially in the HEALTH and EDUCATION domain.	Mandatory Requirement	AM, SSI-HE, HEDF, ATR, STM, DPIA	GDPR, Article 5, 9, 32	

LEG-TSD-004	With regards to the GDPR applicability: TRUSTEE could consider utilising anonymisation tools and provide appropriate explanation of applied anonymisation technique to the end users. However, it should be left up to the data providers whether the data which they would like to share via TRUSTEE should be anonymised or not – sometimes the anonymisation process causes loss of quality of the data set.	Desirable Requirement	HEDF, SSI-HE	GDPR, Article 2, 3, 5(1) (c)	
LEG-TSD-005	TRUSTEE must be equipped with the technical and governance solutions which will allow for appropriate protection of the data which are or may be considered personal data, to ensure that the rights and freedoms of the data subjects (users or people whose data are accessed between users) are respected, can be exercised and will not be infringed.	Mandatory Requirement	AM, SSI-HE, HEDF, ATR, STM, DPIA	GDPR, Article 5, 24 - 30, 32-36, 89 (1)	
LEG-TSD-006	Any website, application or server must incorporate basic state-of-the-art security rules, not only on network communications but also on authentication and infrastructure.	Mandatory Requirement	All ABBs	GDPR, Article 32(1) CNIL GDPR developer guide	
LEG-TSD-007	Security solutions recommended to be considered: - Backups, if possible encrypted and checked regularly should be made, to enable restoring the system. - Limitation of the size of the software stack used, - Installation of critical updates without delay, - Automation of a vulnerability watch, - Application of vulnerability detection tools, - Restriction of physical and software access to diagnostic and remote configuration ports, - Protection of the databases made available online, - Secure database management.	Desirable Requirement	All ABBs	CNIL GDPR developer guide	

LEG-TSD-008	TRUSTEE should have its Terms of Use where it will be clearly indicated what role TRUSTEE system developers and providers assume in the context of the GDPR. The Terms of Use shall be presented to the end users who, if they would like to use tools offered by TRUSTEE to exchange and explore the data which may be considered (encrypted/pseudonymised) personal data, must accept them to be able to do so.	Mandatory Requirement	STM, DPIA	GDPR, Article 24-28	
LEG-TSD-009	In case TRUSTEE system developers decide to engage an external cloud service provider, they shall inform about it the end users and include this information in the Terms of Use of the TRUSTEE system. The data controller (i.e. Data provider / Data consumer), by approving Terms of Use, will have a chance to authorise this sub-processing relationship.	Mandatory Requirement	STM, DPIA	GDPR, Article 24-28	
LEG-TSD-010	When engaging external cloud service provider, the choice must be made according to the security knowledge and the expected quality of service.	Mandatory Requirement	N/A	GDPR, Article 24 and 28	
LEG-TSD-011	TRUSTEE architecture should allow the data provider to indicate what data can be processed, for how long and for what purpose.	Mandatory Requirement	STM, DPIA	GDPR, Article 5, 6	
LEG-TSD-012	With regard to the user data processed on TRUSTEE, TRUSTEE providers should implement solutions, such as technical and organisational measures (such as, in particular, encryption, pseudonymisation, privacy policy, and Terms of Use) which will address the requirements defined in the GDPR principles.	Mandatory Requirement	AM, SSI-HE, HEDF, ATR, STM, DPIA, KR	GDPR, Article 5, 32, 89	

LEG-TSD-013	With regards to the TRUSTEE user's identification data, TRUSTEE providers will be controllers and below principles should be addressed by them: • Lawfulness: A legal basis in Article 6 (and where applicable an exception defined in Article 9 (2) GDPR for special categories of data) should be indicated. • Transparency: Information to the data subject (Data provider and Data consumer) about processing of their data by TRUSTEE Developers and maintaining parties should be presented. • Purpose limitation: Specific, explicit and legitimate purposes and process only data for this set purpose shall be defined. • Data Minimisation: Only data which is relevant and strictly necessary to provide the service can be processed. Pseudonymisation and/or anonymisation of data should be considered wherever possible. In the case of TRUSTEE, SSI authentication should allow for protection of users' privacy. Minimisation of the amount of data collected also in the log data and not storing sensitive or critical data must be applied. • Accuracy: Technical and organisational measures shall be put in place to ensure that the data provided by users is up to date and accurate and/or allow the users to correct inaccurate data. • Storage limitation: All data which is no longer necessary for the fulfilment of the purpose shall be deleted or anonymised, unless an exception applies (e.g., scientific research). Retention periods should be associated for each category of data (including logs). • Confidentiality, Integrity and Data Security: Technical and organisation measures, described in the sections below and in the technical requirements shall be implemented to mitigate risks and to ensure a high level of security, confidentiality and integrity of the data.	Mandatory Requirement	AM, SSI-HE, HEDF, ATR, STM, DPIA, KR	GDPR, Article 5	

LEG-TSD-014	With regards to the data provided onto TRUSTEE by data providers, which may be considered personal, the system should allow to address these principles as well, in particular: • Lawfulness: A legal basis in Article 6 (and where applicable Article 9 GDPR) should be indicated by the data provider as the basis for providing the data onto TRUSTEE and sharing them with the data consumers, and by the Data consumer for further processing for Data consumer's purposes. • Transparency: Data provider and Data consumer should be indicated as responsible for providing data subjects with the information about processing of their data on TRUSTEE and informed that the data consumers are among recipient of the data. • Purpose limitation: Data providers should indicate limitations which are in place in the context of the purposes for which the data provided onto TRUSTEE can be processed. • Data Minimisation: Only data which is relevant to achieve the defined purpose can be processed. Pseudonymisation and/or anonymisation of data should be considered wherever possible. In the case of TRUSTEE, homomorphic encryption is one of the means which should contribute to data providers' and data consumers' compliance with this requirement. • Accuracy: The end users should make sure that the technical and organisational measures shall be put in place on their sides as well as on TRUSTEE allow to keep the data accurate. The users providing the data should confirm whether and how the data provided onto TRUSTEE are accurate, by providing this information in the contract management layer and in the privacy impact assessment. • Storage limitation: All data which is no longer necessary for the fulfilment of the legitimate purpose set out by the Data providers or Data consumers shall be deleted or anonymised, unless an exception applies (e.g., scientific research). • Confidentiality, Integrity and Data Security: Technical and organisation measures, described in the sections below and in the technical requirements shall be implemented to mitigate risks to data processing and to ensure a high level of security, confidentiality and integrity of the data.	Mandatory Requirement	AM, SSI-HE, HEDF, ATR, STM, DPIA, KR	GDPR, Article 5	

LEG-TSD-015	TRUSTEE, as data processor, must allow for responding in a timely manner to data subjects rights by the Data provider (a party sharing the data being personal data with TRUSTEE system providers) to: • Information about the data processing (Article 12-14) • Access (Article 15) • Rectification (Article 16) • Erasure (Article 17) • Restriction on processing (Article 18) • Portability (Article 20) • Object (Article 21) • Not to be subject to a decision based solely on automated processing, including profiling (Article 22).	Mandatory Requirement	STM, DPIA	GDPR, Articles 12-22 and 28	
LEG-TSD-016	TRUSTEE, as data controller of the TRUSTEE user's identification data must respond in a timely manner to data subjects rights to: • Information about the data processing (Article 12-14) • Access (Article 15) • Rectification (Article 16) • Erasure (Article 17) • Restriction on processing (Article 18) • Portability (Article 20) • Object (Article 21) • Not to be subject to a decision based solely on automated processing, including profiling (Article 22) • Lodge a complaint with the supervisory authority (Article 13).	Mandatory Requirement	STM ATR, HE-SSI, DPIA	GDPR, Articles 12-22 and 24	
LEG-TSD-017	The information provided to the user of the platform, being a natural person, whose credentials/log in details will be processed on TRUSTEE, shall fulfil the below requirements: • concise, transparent, intelligible and easily accessible • a clear and plain language • in writing including by electronic means • recommendation to use layered privacy statements. No information required if it proves impossible or would involve a disproportionate effort to contact the data subject.	Mandatory Requirement	DPIA, STM, DA	GDPR, Articles 13 and 14	
LEG-TSD-018	TRUSTEE system developers should represent and describe how the products work, with diagrams presenting data flows and description of processes carried out, in a manner understandable for the users.	Mandatory Requirement	In the project phase: deliverables of TRUSTEE project After the project: a new manner would need to be	GDPR, Article 13 and 14	

			considered by the parties exploiting the solution.		
LEG-TSD-019	TRUSTEE must incorporate the solutions compliant with data protection requirements already in planning and development phases (privacy by design), and ensure that only personal data necessary for the purpose of providing TRUSTEE to the users are processed (privacy by default).	Mandatory Requirement	All ABBs	GDPR, Article 25	
LEG-TSD-020	TRUSTEE must require from the end users only provision of the data necessary to fulfil the purpose of processing (i.e. authentication and necessary access information).	Mandatory Requirement	AM, SSI-HE	GDPR, Article 5(1)(c)	
LEG-TSD-021	The measures to be provided and which should be considered in the process of development and maintenance of TRUSTEE are: • Pseudonymisation and encryption of personal data – the TRUSTEE platform will employ homomorphic encryption technique; • Measures to ensure the security of IT employed in the data processing, i.e. ensuring its ongoing confidentiality, integrity, availability and resilience – to answer these requirements, TRUSTEE will employ authentication services, such as SSI and SDK; • Incident response capabilities to restore the availability and access to personal data in a timely manner – due to the federated learning implementation, this responsibility would be rather dependent on the end-users providing datasets via TRUSTEE; • A regular assessment and evaluation of the effectiveness of the implemented technical and organisational security measures. The list is non-exhaustive. The TRUSTEE providers are encouraged to provide additional security measures to ensure appropriate level of data security.	Mandatory Requirement	All ABBs	GDPR, Article 32 (1)	
LEG-TSD-022	Appropriate, secure and differential data access management policies must be applied. Authentication should be imposed before any access to personal data.	Mandatory Requirement	AM, SSI-HE, STM, DPIA	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-023	Logging systems should be used to detect anomalies or security events.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-024	Penetration testing should be considered and performed.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-025	Efficient version control system should be set up, to track different versions of the code over the time, with appropriate authentication and permission system and regular backups, as well as efficient development system.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	

LEG-TSD-026	Code quality metrics tools should be considered.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-027	Control of the use libraries and SDKs should be applied.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-028	When using free or open source software, TRUSTEE system developers should choose solutions that have an active community, good documentation and are regularly updated.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-029	When using types of solutions with commercial support, TRUSTEE system providers should contractually ensure maintenance and updates of the code during the whole lifecycle of the project.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-030	Change of the default configurations, audit, mapping of the dependencies of the SDKs and third party libraries should be considered.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-031	Implementation of cryptographic algorithms and protocols by TRUSTEE system developers by their own should be avoided, and choosing cryptographic libraries which are maintained, easy to use and recognised should be encouraged.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-032	Appropriate maintenance of libraries and SDKs should be considered, e.g. using dependency management system, managing updates to the dependencies, monitoring the versions of libraries and SDKs and regularly check the open source projects and their status, especially in the context of the change of domain or package ownership.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-033	Quality and security of the code and its documentation, as well as the documentation of the system architecture should be ensured.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-034	Applications must be tested (development test, security test). Real/ production data should not be used during the development and testing phase, as using personal data for testing purposes may go beyond their initial purpose.	Desirable Requirement	All ABBs	GDPR, Article 32 (1), CNIL GDPR developer guide	
LEG-TSD-035	When data security is compromised, TRUSTEE system providers must report the violation to the data controllers (Data providers and Data consumers) without undue delay if it is likely to pose a risk to the rights and freedoms of individuals.	Mandatory Requirement	All ABBs	GDPR, Article 28 and 33(2)	

LEG-TSD-036	In case of the transfer of the data via TRUSTEE to third (non-EU) countries, appropriate safeguards must be considered and TRUSTEE should enable their implementation by the user.	Mandatory Requirement	AM, SSI-HE, OneSS, ATR, DPIA	GDPR, Article 46	
LEG-TSD-037	TRUSTEE should allow the users to conduct a DPIA prior to any data processing activity (i.e. prior to accessing (encrypted) data and conducting computing activities with the use of data).	Mandatory Requirement	DA, STM, DPIA	GDPR, Article 35	
REQUIREMENTS APPLICABLE TO PERSONAL AND NON-PERSONAL DATA					
LEG-TSD-038	TRUSTEE should provide easy data portability to their professional users.	Mandatory Requirement	to be further investigated	Regulation on free flow of non-personal data	
LEG-TSD-039	Implement measures into the solutions of TRUSTEE for the protection of personal data in mixed data sets that are to be brought by Data providers.	Mandatory Requirement	DPIA, to be further investigated	Regulation on free flow of non-personal data	
LEG-TSD-040	Evaluate the adoption of the codes of conduct for Iaas and SaaS developed under the Regulation on Free Flow of Non-Personal Data for the porting of data brought by business users.	Optional Requirement	to be further investigated	Regulation on free flow of non-personal data	
LEG-TSD-041	Explore whether any specific addressing of tools/considerations of allowing for the sharing of data/documents held by public sector bodies could be facilitated by TRUSTEE.	Mandatory Requirement	All ABBs	Open Data Directive	
LEG-TSD-042	TRUSTEE system developers should identify and take appropriate and proportionate technical and organizational measures to manage risks created by the security systems.	Mandatory Requirement	All ABBs	NIS Directive	

LEG-TSD-043	The level of security which should be ascertained has to be appropriate to the risks posed, and should take into account the following elements: a. the security of systems and facilities; b. incident handling; c. business continuity management; d. monitoring, auditing and testing; and d. compliance with international standards.	Mandatory Requirement	All ABBs	NIS Directive, Article 16	
LEG-TSD-044	The security measures taken by TRUSTEE system developers shall include: (a) detection processes and procedures maintained and tested to ensure timely and adequate awareness of anomalous events; (b) processes and policies on reporting incidents and identifying weaknesses and vulnerabilities in their information systems; (c) a response in accordance with established procedures and reporting the results of the measure taken; (d) an assessment of the incident's severity, documenting knowledge from incident analysis and collection of relevant information which may serve as evidence and support a continuous improvement process.	Mandatory Requirement	All ABBs	NIS Directive; Article 2(1) Implementing Regulation	
LEG-TSD-045	TRUSTEE system developers must fulfil the obligation of monitoring, auditing and testing, there should be policies implemented which will define supervision, inspection and verification procedures.	Mandatory Requirement	All ABBs	NIS Directive; Article 2(4) Implementing Regulation	
LEG-TSD-046	TRUSTEE providers may be recognised either as essential (as provider of trusted services) or as important (as research institutions) entities, therefore they should adhere, in all phases of the development of TRUSTEE and its further maintenance, to the NIS 2 Directive stemming obligations, which will be further transposed and established at the national level.	Mandatory Requirement	All ABBs	NIS 2 Directive, Article 2	Law in force, but needs to be transposed by the Member States

LEG-TSD-047	TRUSTEE system developers should consider and implement the requirements stemming from the NIS and NIS 2 Directive, to ensure appropriate level of security of information, especially implementation of the means such as: - policies on risk analysis and information system security, handling, incident - business continuity - supply chain security, - security in acquisition, development and maintenance of network and information system, - assessment of the effectiveness of implemented measures, - cybersecurity training and basic cyber hygiene practices, - procedures regarding the use of cryptography and, where appropriate, encryption, access control policies, asset management and HR security, - the use of multi-factor authentication or continuous authentication solutions within the entity, where appropriate.	Mandatory Requirement	All ABBs	NIS 2 Directive, Article 21 (2)	
LEG-TSD-048	TRUSTEE system developers shall implement appropriate incident reporting procedures.	Mandatory Requirement	All ABBs	NIS 2 Directive, Article 23	
LEG-TSD-049	TRUSTEE providers shall disclose the terms and conditions that apply to the services provided by them, as well as shall clearly describe the policies governing their provision.	Mandatory Requirement	All ABBs	NIS 2 Directive, Article 21	
LEG-TSD-050	TRUSTEE providers shall thoroughly plan and implement appropriate security measures and constantly monitor, test and audit them, as well as keep the documentation describing those measures in a transparent and clear manner.	Mandatory Requirement	All ABBs	NIS 2 Directive	
LEG-TSD-051	TRUSTEE system developers need to understand eIDAS framework, as they aim to build upon the eIDAS trust framework and ensure full compatibility with it in developing a self-sovereign identity management.	Mandatory Requirement	AM, SSI-HE	eIDAS	
LEG-TSD-052	In case electronic signature is applied for the log-in, the e-Signature should fulfil the requirements indicated in the Regulation.	Mandatory Requirement	AM, SSI-HE	eIDAS	Law and/or guideline in force and applicable
LEG-TSD-053	TRUSTEE shall fulfil the requirements imposed on the trust services and TRUSTEE system developers and Partners responsible for provision of the authentication systems should fulfil the obligations listed in Chapter III of the eIDAS.	Mandatory Requirement	AM, SSI-HE	eIDAS	
LEG-TSD-054	Identity and access management using SSI and eID (blockchain- based IAM) must support GDPR compliance, asserted and guaranteed by service provider.	Mandatory Requirement	AM, SSI-HE	eIDAS, GDPR, Kuperberg 2019 (publication)	

LEG-TSD-055	SSI over eIDAS must enable control over geographical distribution of data (restriction to the EU or to specific countries should be possible).	Mandatory Requirement	AM, SSI-HE	eIDAS, GDPR, Kuperberg 2019 (publication)	
LEG-TSD-056	Access information/material and credentials must be possible to be withdrawn or rotation.	Mandatory Requirement	AM, SSI-HE	eIDAS, GDPR, Kuperberg 2019 (publication)	
LEG-TSD-057	Loggs, identity changes and access propperties and rights must be GDPR compliant.	Mandatory Requirement	AM, SSI-HE	eIDAS, GDPR, Kuperberg 2019 (publication)	
LEG-TSD-058	End user (Data provider, Developer and Data consumer) data cannot be shared by default, end-users should have the possibility to opt-in and further opt-out, if they want to. Data cannot be evaluated or sold by the identity provider.	Mandatory Requirement	AM, SSI-HE	eIDAS, GDPR, Kuperberg 2019 (publication)	
LEG-TSD-059	Identity theft must be prevented by applying control of identity export and transfer.	Mandatory Requirement	AM, SSI-HE	eIDAS, GDPR, Kuperberg 2019 (publication)	
LEG-TSD-060	Hosting model must be selected (provider-offered service / customer-managed on-premise hosting / provider- managed cloud hosting / customer-managed cloud hosting)	Mandatory Requirement	AM, SSI-HE	eIDAS, Kuperberg 2019 (publication)	
LEG-TSD-061	Usage model of SSI over eIDAS should include rights to audit or to re-use the source code - reuse rights may be limited (e.g. as with GPL open-source license).	Desirable Requirement	AM, SSI-HE	eIDAS, Kuperberg 2019 (publication)	
LEG-TSD-062	IAM network traffic to and from IAM-using backend applications should be separated from Internet traffic.	Desirable Requirement	AM, SSI-HE	eIDAS, Kuperberg 2019 (publication)	
LEG-TSD-063	TRUSTEE should investigate whether TRUSTEE solution as a whole is considered an intermediate service with the goal of establishing commercial relationships for data sharing between an undetermined number of data holders (Data providers) on one hand and data users (TRUSTEE - Data consumers) on the other, through technical, legal, or other means, including for data subjects to exercise their rights in relation to personal data.	Desirable Requirement	The outcome will have an impact on all ABBs	DGA, Article 1(b) and Article 2(11)	Law in force, but applicable from 24 September 2023

LEG-TSD-064	If TRUSTEE is considered data intermediation service, TRUSTEE system providers must further determine whether TRUSTEE will be an intermediation service between data holders and potential data users, including making available the technical or other means to enable such services, that would include bilateral or multilateral exchanges of data or the creation of platform enabling the exchange or joint use of data, as well as the establishment of other specific infrastructure for the interconnection of data holders (Data providers) with data users (Data consumers).	Desirable Requirement	The outcome will have an impact on all ABBs	DGA, Article 10(a)	
LEG-TSD-065	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE service providers must submit a notification to the competent authority for data intermediation services before starting their activity.	Desirable Requirement	This is for exploitation phase.	DGA, Article 11(1)	
LEG-TSD-066	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE service providers cannot use the data for which it provides data intermediation services for purposes other than to put them at the disposal of data users (Data consumers).	Desirable Requirement	SSI-HE, HEDF, ATR, KR, DL, AIMaaS, to be addressed also in ToU of TRUSTEE	DGA, Article 12(a)	
LEG-TSD-067	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE service providers must not make the use of TRUSTEE dependent on the use of their other services.	Desirable Requirement	ALL ABBs - to be addressed also in ToU of TRUSTEE	DGA, Article 12(b)	
LEG-TSD-068	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE service providers must use the data of the data holders (Data providers/Model providers) and data users (Data consumers) and their activities only for the development of the service and they should be informed about it.	Desirable Requirement	DA, STM - to be addressed in ToU of TRUSTEE	DGA, Article 12 (c)	
LEG-TSD-069	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE service providers must facilitate exchange of the data in the format they received them and convert the data only to enhance interoperability or if requested by the data user or where required by EU law. TRUSTEE service providers should offer opt-out regarding these conversions to Data Providers (data holders).	Desirable Requirement	SSI-HE, HEDF, DL, KR	DGA, Article 12(d)	

LEG-TSD-070	If TRUSTEE happens to be intermediation service according to the DGA, TRUSTEE service providers may offer additional services to data holders or data subjects that facilitate the exchange of data, such as temporary storage, curation, conversion, anonymisation and pseudonymisation; such tools should be used only at the explicit request or approval of the Data Provider (data holder) or data subject and third-party tools offered in that context should not be used for other purposes.	Desirable Requirement	ALL ABBs, to be also indicated in ToU of TRUSTEE	DGA, Article 12(e)	
LEG-TSD-071	If TRUSTEE happens to be intermediation service according to the DGA, TRUSTEE service providers must provide fair, transparent and non-discriminatory access procedure (including prices and terms of service).	Desirable Requirement	AM, SSI-HE, STM, DPIA, DA - also to be reflected in ToU of TRUSTEE	DGA, Article 12(f)	
LEG-TSD-072	If TRUSTEE happens to be intermediation service according to the DGA, TRUSTEE service providers must have procedures against fraudulent or abusive practices concerning access to the data via services.	Desirable Requirement	AM, STM, ATR, TAI-SDF, DPIA, to be also considered in ToU of TRUSTEE	DGA, Article 12(g)	
LEG-TSD-073	If TRUSTEE happens to be intermediation service according to the DGA, TRUSTEE service providers must ensure continuity of the services, and where storage of data is included, allow parties using services to access, transfer or retrieve their data.	Desirable Requirement	SSI-HE, AM, KR, DL, HEDF - to be also considered in ToU of TRUSTEE	DGA, Article 12(h)	
LEG-TSD-074	If TRUSTEE happens to be intermediation service according to the DGA, TRUSTEE service providers must ensure interoperability by adhering to common open standards in the sectors where service operates.	Desirable Requirement	SSI-HE, KR, AIMaaS, DL, HEDF	DGA, Article 12(i)	
LEG-TSD-075	If TRUSTEE happens to be intermediation service according to the DGA, TRUSTEE service providers must apply measures preventing transfer of or access to non-personal data that is unlawful under EU or national law.	Desirable Requirement	AM, SSI-HE, DL, ATR, STM	DGA, Article 12(j)	
LEG-TSD-076	If TRUSTEE happens to be intermediation service according to the DGA, TRUSTEE service providers must inform data holders (Data providers) in the event of an unauthorised transfer, access or use of the non-personal data that they have shared.	Desirable Requirement	Could be addressed by Terms of Use of TRUSTEE, as an obligation for the service provider or by virtual assistant in the DA	DGA, Article 12(k)	

LEG-TSD-077	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE service providers must take necessary measures to ensure an appropriate level of security for the (...) processing and transmission of non-personal data and the highest level of security for transmission of competitively sensitive information (e.g. trade secrets).	Desirable Requirement	ALL ABBs	DGA, Article 12(l)	
LEG-TSD-078	If TRUSTEE happens to be intermediation service according to the DGA , and where TRUSTEE system providers provide tools for obtaining permissions to process data made available by Data providers (data holders), TRUSTEE system providers must specify non-EU jurisdictions in which the data use is intended to take place and provide tools to the Data provider to give and/or withdraw permission to process data.	Desirable Requirement	AM, STM	DGA, Article 12(n)	
LEG-TSD-079	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE system providers must maintain a log record of the intermediation activities.	Desirable Requirement	ATR, STM, to be further investigated	DGA, Article 12(o)	
LEG-TSD-080	If TRUSTEE happens to be intermediation service according to the DGA , TRUSTEE system providers must apply TLOMs (technical, legal, organisational measures), including contractual arrangements, to prevent international transfer or governmental access to non-personal data in the EU where it would not comply with the law (unless based on an international agreement between the countries or other exceptions from DGA apply) , and should inform Data provider about the request to access the data except where it serves law enforcement.	Desirable Requirement	AM	DGA, Article 31	

LEG-TSD-081	If TRUSTEE becomes a data processing service that is not free of charge and does not operate on a trial basis, TRUSTEE system developers should enable the users to change to another service and shall not impose and shall remove commercial, technical, contractual and organisational obstacles for that, defined in Article 23(1).	Optional Requirement	SSI-HE, KR, ATR, to be further investigated	Data Act, Article 23(1) and Article 26a(2) (EP version)	<u>Legal Act at the proposal stage, not adopted, subject to changes !</u>
LEG-TSD-082	If TRUSTEE becomes a data processing service that is not free of charge and does not operate on a trial basis, the contracts between customers and providers of data processing services shall include: a) clauses allowing to switch to a data processing service offered by another provider or to port all data, applications and digital assets generated by the customer on an on-premise system; b) an exhaustive specification of all data and application categories exportable during the switching process, at minimum – all data imported by the customer and all data and metadata created by the customer and by the use of service (configuration parameters, security settings, access rights and logs); c) a minimum period of data retrieval of at least 30 calendar days; d) an obligation on the provider of data processing services to delete all of the former customer’s exportable data after the expiration of the minimum period from point c).	Optional Requirement	SSI-HE, KR, OneSS (External Connectors), DL, ATR, to be further investigated	Data Act, Article 24 and Article 26a(2) (EP version)	
LEG-TSD-083	If TRUSTEE becomes a data processing service that is not free of charge and does not operate on a trial basis, TRUSTEE shall make open interfaces publicly available and free of charge and ensure compatibility with open interoperability specifications of European standards for interoperability. If such standards do not exist, the data should be provided in a structured, commonly used and machine-readable format.	Optional Requirement	SSI-HE, KR, OneSS (External Connectors), DL, ATR, to be further investigated	Data Act, Article 26 (2) and Article 26a(2) (EP version)	
LEG-TSD-084	Where TRUSTEE providers see that the transfer of or access to non-personal data can cause the risk of re-identification of non-personal or anonymised data, they should request authorisation from data protection authorities.	Desirable Requirement	to be investigated	Data Act, Article 27(4a) (EP version)	

LEG-TSD-085	TRUSTEE system providers (as parties offering data services to other participants) shall ensure interoperability, present at the level of data, data sharing mechanisms and services (including the technical means to do so): a) the dataset content, use restrictions, licences, data collection methodology, data quality and uncertainty shall be sufficiently described in a machine-readable format to allow the recipient to find, access and use the data; b) the data structures, data formats, vocabularies, classification schemes, taxonomies and code lists shall be described in a publicly available and consistent manner; c) the technical means to access the data, such as application programming interfaces, and their terms of use and quality of service shall be sufficiently described to enable automatic access and transmission of data between parties, including continuously or in real-time in a machine-readable format; d) the means to enable the interoperability of contracts for data sharing within their services and activities shall be provided.	Desirable Requirement	SSI-HE, DL, HEDF, STM	Data Act, Article 28 (1) (EP version)	
LEG-TSD-086	TRUSTEE system providers, as the group being one of the participants of data spaces, must agree on the rules by which the accountabilities regarding those requirements from LEG-TSD-085 are defined between the participants (Data providers, Data Consumers, Model Providers, Developers).	Desirable Requirement	Could be addressed by Terms of Use of TRUSTEE	Data Act, 28(3a) (EP version)	
LEG-TSD-087	TRUSTEE system developers, by providing the service which involves deployment of smart contracts for others in the context of an agreement to make data available, shall ensure that the smart contract has been designed to offer a very high degree of robustness to avoid functional errors and to withstand manipulation by third parties (robustness).	Desirable Requirement	ATR, STM, DPIA	Data Act, Article 30(a) (EP version)	
LEG-TSD-088	TRUSTEE system developers should ensure that a mechanism exists to terminate the continued execution of transactions: the smart contract shall include internal functions which can reset or instruct the contract to stop or interrupt the operation to avoid future (accidental) executions (safe termination and interruption). The conditions under which a smart contract could be reset or instructed to stop or interrupted, should be clearly and transparently defined. Especially, it should be assessed under which conditions non-consensual termination or interruption should be permissible.	Desirable Requirement	ATR, STM, DPIA	Data Act, Article 30(b) (EP version)	
LEG-TSD-089	TRUSTEE system developers should ensure that a smart contract is protected through and designed to offer rigorous access control mechanisms at the governance and smart contract layers (access control).	Desirable Requirement	ATR, STM, DPIA	Data Act, Article 30(a) (EP version)	
LEG-TSD-090	TRUSTEE system developers should ensure that a smart contract affords the same level of protection and certainty as any other contracts generated through different means (equivalence).	Desirable Requirement	ATR, STM, DPIA	Data Act, Article 30(ba) (EP version)	

LEG-TSD-091	TRUSTEE system developers should ensure that a smart contract ensures the confidentiality of trade secrets.	Desirable Requirement	ATR, STM, DPIA	Data Act, Article 30(bb) (EP version)	
LEG-TSD-092	TRUSTEE system developers should determine whether, by allowing sharing the AI solutions by the Model providers (users providing AI technology to Data providers and Data consumers), they also participate in providing AI systems. If so, AI Act requirements will apply in the future.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 3(2), Article 28 (Czech Presidency version)	
LEG-TSD-093	If TRUSTEE system developers are considered AI system providers , TRUSTEE system developers must determine if the AI is a high risk AI .	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 6(1) and (2), Article 6(3) and Annex III, Article 28	
LEG-TSD-094	If TRUSTEE system developers occur to be providers of high risk AI , a risk management system must be established, implemented, documented and maintained. Further details are included in the AIA.	Desirable Requirement	KR, ATR, TAI-SDF, DPIA, AIMaaS,	AIA, Article 9	
LEG-TSD-095	If TRUSTEE system developers occur to be providers of high risk AI , the training, validation and testing data sets must meet a certain quality criteria from Article 10, if the high-risk AI system uses techniques including the training of models with data. The data governance practices must concern in particular: relevant design choices, transparency as regards the original purpose of data collection, data collection process, data preparation operations (annotation, labelling, cleaning, updating, enrichment and aggregation), the formulation of assumptions, an assessment of the availability, quantity and suitability of the needed datasets, examination in view of possible biases, appropriate measures to detect, prevent and mitigate possible biases, identification of data gaps or shortcomings and how they can be addressed. At the level of individual datasets and their combination, the training, validation and testing datasets (incl. labels) shall be relevant, representative, appropriately vetted for errors and complete, have appropriate statistical properties, including, as regards the persons or groups of persons in relation to whom the high-risk AI system is intended to be used.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 10(2)-(5)	
LEG-TSD-096	If TRUSTEE system developers occur to be providers of high risk AI , the events in the AI system must be automatically recorded (logs must be kept), to enable tracing of AI system's functionality and to monitor the situations posing risks.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 12, Article 20 and Annex IV	
LEG-TSD-097	If TRUSTEE system developers occur to be providers of high risk AI , a technical documentation of the high-risk AI system must be provided before it is placed on the market or put into service, and regularly updated.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 11	

LEG-TSD-098	If TRUSTEE system developers occur to be providers of high risk AI, TRUSTEE system developers providing AI systems must ensure that their (AI systems') operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately, by providing instructions in an appropriate digital format and information about the purpose, capabilities and limitations of the AI system in a clear manner. Specific parts of the information are included in Article 13 AIA. Additionally, the AI systems must be designed and developed in a way that they can be effectively overseen by natural persons during the time the AI system is in use. This refers also to appropriate human-machine interface tools.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 13, Article 14 and Article 52	
LEG-TSD-099	If TRUSTEE system developers occur to be providers of high risk AI, shall ensure appropriate level of accuracy, robustness, cybersecurity. The levels of accuracy and accuracy metrics shall be declared in the instructions of use. The AI systems must be adaptable against errors, faults and inconsistencies that may occur within the system or the environment in which it operates and withstand any attempts by unauthorised parties to alter the use or performance of the AI system.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 15	
LEG-TSD-100	If TRUSTEE system developers occur to be providers of high risk AI, TRUSTEE system developers must establish and document a post-monitoring plan for high-risk AI systems to actively and systematically collect, document and analyse data provided by users or other data collected on the performance of the AI system throughout their lifecycle	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 61	
LEG-TSD-101	If TRUSTEE system developers occur to be providers of high risk AI, TRUSTEE system developers must provide appropriate quality management system and implement a conformity assessment procedure for the AI system before it is placed on the market.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 16 and Article 17	
LEG-TSD-102	If TRUSTEE system developers occur to be providers of high risk AI, TRUSTEE system developers must communicate clearly that a user being a natural person interacts with an AI system, unless evident from the context.	Desirable Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	AIA, Article 52	

LEG-TSD-103	If TRUSTEE system developers occur to be providers of high risk AI , TRUSTEE system developers must adhere to the requirements for Trustworthy AI: Human Agency and Oversight, Technical Robustness and Safety, Privacy and Data Governance, Transparency, Diversity, Non-discrimination and Fairness, Environmental and Societal Well-Being, Accountability, further described in the legal framework and in socio-ethical requirements section.	Mandatory Requirement	ATR, TAI-SDF, DPIA, AIMaaS, KR	EC Ethics guidelines on trustworthy AI, Assessment List TAI	Standards in force and applicable to TRUSTEE
LEG-TSD-104	If TRUSTEE is considered as an online intermediary service , TRUSTEE system developers and providers must cooperate with national authorities upon the receipt of an order to act against illegal content or to provide specific information about service recipients.	Mandatory Requirement	ALL ABBs	DSA, Article 9 and 10.	Law and/or guideline in force and applicable
LEG-TSD-105	If TRUSTEE is considered as an online intermediary service , TRUSTEE system developers and consortium should designate points of contact both for the recipients of their services and for member states' authorities, the European Commission and the European Board for Digital Services	Mandatory Requirement	ALL ABBs	DSA, Article 11 and 12.	
LEG-TSD-106	If TRUSTEE is considered as an online intermediary service , TRUSTEE should inform their users in TRUSTEE Terms of Use of “any restrictions that they impose in relation to the use of their service in respect of information provided by the recipients of the service” in a “clear, plain, intelligible, user friendly and unambiguous language” and be “in an easily accessible and machine-readable format”. Should there be any significant changes to these terms and conditions, the TRUSTEE service providers/ system developers need to notify their data providers.	Mandatory Requirement	DA, STM	DSA, Article 14.	
LEG-TSD-107	TRUSTEE system developers must abstain from designing and operating their interfaces in a manipulative or deceitful manner or in a way hindering service recipients' ability to make independent and informed choices. (This is an additional requirement for providers of online platforms, under the category of hosting services. To the extent that TRUSTEE's activities and service provision falls under this category, this could be applied.)	Mandatory Requirement	DA, STM, SSI-HE, DPIA, AIMaaS, to be further investigated	DSA, Article 25.	
LEG-TSD-108	TRUSTEE should aim to transpose transparency obligations for contracts between Data providers/Data consumers being business entities and TRUSTEE service providers, as identified in the legal framework.	Desirable Requirement	STM	P2B Regulation	

LEG-TSD-109	TRUSTEE system developers should use only such blockchain solutions which allow to address GDPR principles, especially data minimisation, storage limitation and where necessary, the ones which will allow deletion of the data, or full anonymisation of the data that are not necessary or which erasure was requested by a data subject or a Data provider.	Mandatory Requirement	ATR, SSI-HE, STM, DL	GDPR and its impact on blockchain	Law and/or guideline in force and applicable
LEG-TSD-110	When personal data are stored on an immutable blockchain, the TRUSTEE system developers should ensure one of the points a-d: a) the data are either not or will no longer constitute personal data, b) that data subject (user) is considered to be the controllers of their data, c) that the household exemption defined in the GDPR applies or d) a justification for constant storage of data on the blockchain has to exist.	Mandatory Requirement	HE-SSI, ATR, STM, HEDF	GDPR and its impact on blockchain, G. Kondova, and J. Erbguth. (2020)	
LEG-TSD-111	The contracts developed by the TRUSTEE system providers should be transparent, taking into account the rights to data and information of all parties involved, respect commercial intrests of the parties, ensure undistorted competition, enable data portability as much as possible, define the conditions for access to the data, possibility and limitations concerning re-use of the data, indicate technical measures for data access and sharing, clarify liability towards the parties and indicate parties' rights.	Mandatory Requirement	STM, ATR, DA, TAI-SDF, DPIA	Data Act, DGA, Commission Staff Working Document (2018)	
LEG-TSD-112	TRUSTEE should allow for concluding agreements (e.g. confidentiality agreements) where Data provider could indicate limitations for data processing stemming from the intellectual property rights and trade secrets regulations.	Mandatory Requirement	STM, AM, ATR, OneSS	Infosoc Directive, Software Directive, Database Directive, Trade Secrets Directive	
LEG-TSD-113	TRUSTEE should allow for concluding (licensing) agreements where Model Provider could indicate terms and limitations to the use of AI/ML models provided, on which they assert control stemming from intellectual property rights and/or contracts.	Mandatory Requirement	STM, ATR, DA, DL	Infosoc Directive, Software Directive, Database Directive, Trade Secrets Directive	
LEG-TSD-114	TRUSTEE providers must implement all reasonable technical, legal and organisational measures, including contracts, to prevent international transfer and third-country governmental access to non-personal data held in the UE , where that transfer would cause a conflict with EU or national law (with exceptions in Article 27(2) and (3) Data Act).	Desirable Requirement	STM, could be also addressed by Terms of Use of TRUSTEE	Data Act, Article 27(1)-27(4) and 27(5) (EP version)	<u>Legal Act at the proposal stage, not adopted, subject to changes !</u>

LEG-TSD-115	TRUSTEE system developers should implement solutions allowing to foresee, if a smart contract must be terminated or deactivated, a possibility to archive transactional data, the smart contract logic and code to keep the record of the operations performed on the data in the past (auditability) (data archiving and continuity).	Desirable Requirement	ATR, STM, DPIA	Data Act, Article 30 (EC proposal version)	

ANNEX 6: SOCIO-ETHICAL REQUIREMENTS

6.1 SOCIO-ETHICAL REQUIREMENTS FOR TRUSTEE SYSTEM DEVELOPERS¹⁰³⁷

Nature of requirement	Req ID	Requirement (ethical principles and ethical imperatives) (Technical, Management, Legal teams + AI Model Providers)	Priority (Must-have. Mandatory Requirement, Should-have. Desirable Requirement, Could-have. Optional Requirement, Will-not-have. Possible Future Enhancement)	Legal Act/ Relevant Guideline/Standard	No of related legal requirements for TRUSTEE system developers and end users

¹⁰³⁷ *points relating to mapping (from about LEG-USR-055 to about LEG-USR-068, LEG-USR-077 and LEG-USR-078)
any aspects relating to legal req for end user and re-use of data
in relation to TRUSTEE use by public sector body
in relation to intellectual property rights (LEG-TSD-112-113 and in LEG-USR-061, 094, 096 and 097)

Adhering to fundamental human rights	SOC-ETH-TSD-001	Respect for democracy, justice, and the rule of law. Any AI deployment must ensure it respects the plurality of choice, values, and decision-making of individuals.	Mandatory	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-002, LEG-TSD-003, LEG-TSD-013, LEG-TSD-014, LEG-TSD-040, LEG-TSD-043, LEG-TSD-047; LEG-USR-002, LEG-USR-025, LEG-USR-036-037 (in relation to fundamental rights), LEG-USR-059, LEG-USR-062, LEG-USR-092
	SOC-ETH-TSD-002	Respect citizens' rights (regardless of circumstances of individuals, e.g., illegal immigrants) in deploying AI onto open data markets such as TRUSTEE	Mandatory	EU and fundamental human rights; EC expert guidelines for trustworthy AI; HRESIA Model; DPIA; ICO Ethics framework (2020) https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020	
	SOC-ETH-TSD-003	Ensure that any AI and data manipulation deployed is done for the purpose* of enhancing and benefiting human and societal wellbeing and human capabilities.	Mandatory	EC expert guidelines for trustworthy AI; AIA; *Any AI deployment should have a clearly defined purpose - principle 1, CNIL (2022)	
	SOC-ETH-TSD-004	Provide safeguards against overconfidence and overreliance on any of the data manipulation AI provisions on TRUSTEE.	Desirable	EC Ethics Guidelines for Trustworthy AI	
	SOC-ETH-TSD-005	Prioritise sustainable development, personal data rights, and agency over digital identity; and ensure legal frameworks for accountability are met.	Desirable	EC Ethics Guidelines for Trustworthy AI	

	SOC-ETH-TSD-006	Ensure evaluation and assessment phases of AI and data manipulation: first to identify if AI is of high risk; secondly to (re)identify bias, direct/indirect prejudice and discrimination against certain groups of people, to update algorithms' programming, re-evaluate and address AI functionalities' purpose, constraints, requirements, and decisions to maintain transparency and prevent harm from bias and exacerbating prejudice and discrimination.	Desirable	AI Now (AI Now: https://ainowinstitute.org/our-work); AIA Article 3; FRA (2018). Big data discrimination in data-supported decision making. Luxembourg, Publications Office. Accessible from: https://fra.europa.eu/en/publication/2018/bigdata-discrimination-data-supported-decision-making	
	SOC-ETH-TSD-007	Ensure universal design for accessibility, explicability and explainability specifically where non-technical end-users engage with open data platforms such as TRUSTEE. Their experience and understanding of AI and data manipulation, their age, abilities, and other characteristics reflecting how use and interpretation of the produced results from TRUSTEE are accepted should be considered by TRUSTEE's developers.	Mandatory	Article 42 of the Public Procurement Directive of the European Parliament on technical specifications around design (e.g., EN301 549); CNIL Principle 8: "providing information and explicability"	
	SOC-ETH-TSD-008	Ensure fairness through transparency around data manipulation and AI use to achieve equal and fair distribution of benefits versus costs ('white box' or 'glass box' design versus 'black box'); advantages versus disadvantages; To ensure fairness, make explicit and identifiable who arrives at the decisions or results (from the data manipulation and AI-enabled data processing) .	Desirable	EC Ethics Guidelines for Trustworthy AI; HRESIA Model	

Privacy-by-design, security-by-design, lawful-by-design for transparency and accountability LEG-TSD-003-007, LEG-TSD-012-014, LEG-TSD-043; LEG-USR-95.	SOC-ETH-TSD-009	Provide traceability, logging of any AI processes and outcomes to enable auditability of any of the processes that underpin data manipulation and AI in concrete contexts and the level of severity of consequences (margins of error or in cases of risks from inaccuracy of data or AI models)	Mandatory	EC Ethics Guidelines for Trustworthy AI; HRESIA Model; AIA; https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682	LEG-TSD-092; LEG-USR-081-092
	SOC-ETH-TSD-010	Provide accountability through transparency: provide clarity about the purpose limitations of AI functionalities, where algorithmic models have been derived and what their limitations are.	Mandatory	EPRS European Parliamentary Research(STOA)	LEG-TSD-092, LEG-TSD-093, LEG-TSD-095.
	SOC-ETH-TSD-011	Ensure that data privacy and protection are guaranteed. Any AI and data manipulation should prior to their work have guaranteed privacy, security and protection throughout the functionalities' and the platform's entire lifecycle.	Mandatory	GDPR	LEG-TSD-107; LEG-USR-026.

	SOC-ETH-TSD-012	Be able to demonstrate how algorithms have been trained - what data input was used, and how this evolved throughout the project's lifecycle.	Desirable	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-092-93 , LEG-TSD-095 ; LEG-USR-016 .
	SOC-ETH-TSD-013	Be able to explain how the platform's AI functionalities and data manipulation may influence end users of TRUSTEE.	Desirable	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-092-093 , LEG-TSD-095 .
	SOC-ETH-TSD-014	Prevent harm through technological security and safety of the systems that handle data and data processing; provide sufficient information around data manipulation and AI (xAI that is transparent and understandable to end users on how manipulation and AI arrive at specific results). Particular attention must be paid in relation to vulnerable groups such as children, persons with disabilities, historical disadvantages, and any other individuals or collectives who are at risk of exclusion.	Mandatory	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-003-006 , LEG-TSD-012 , LEG-USR-024 .
	SOC-ETH-TSD-015	Ensure that there are proper measures to assess risks and safety and clearly record these to develop (communications) protocols, policies and procedures.	Desirable	EC Ethics Guidelines for Trustworthy AI; AIA, Article 6(1) and (2)	LEG-TSD-092-093 , LEG-TSD-095 .

	SOC-ETH-TSD-016	Provide mitigation plan for any foreseeable risks of data manipulation or malicious use of the platform (hack of the data or systems etc).	Optional	EC (2021) Europe fit for the digital age...Accessible from: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682	LEG-TSD-092-93, LEG-TSD-095.
	SOC-ETH-TSD-017	Consider any liabilities relating to consumer protection rules both for data providers, and the end users of TRUSTEE (especially with regards to the use of AI and data manipulation).	Desirable	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-014 LEG-USR-25, LEG-USR-040-042, and any other impact assessment relating points
	SOC-ETH-TSD-018	Consider any risks relating to AI and data manipulation providing wrong results, or the impact of the platform becoming unavailable or providing societally unacceptable results such as discriminatory or biased.	Desirable	European Union Agency for Fundamental Rights 2018	LEG-TSD-035, LEG-TSD-042-043, LEG-TSD-084; LEG-USR-025.
Prevent harm	SOC-ETH-TSD-019	Provide data minimisation wherever legally bound, appropriate and necessary, and provide reporting of negative impact, trade-offs and redress through technical mechanisms. Document any potential trade-offs by identifying relevant interests and values that may be implicated by the AI system. Document any decision that is made on this trade-off (if information is disclosed based on data manipulation and this one benefits one organisation over an individual or another organisation, what decision was made to allow the information to be disclosed through TRUSTEE - such information should be clearly documented).	Desirable	GDPR and EC Ethics Guidelines for Trustworthy AI	LEG-TSD-013-014

	SOC-ETH-TSD-020	Provide reproducibility of results to ensure accuracy of the AI and data manipulation functionalities. AI functionalities should be able to produce the same results in a range of situations. This must be clear and evident for the robustness and credibility of open data platforms such as TRUSTEE.	Desirable	AI Act; EC Ethics Guidelines for Trustworthy AI	LEG-TSD-111-113 LEG-USR-042 LEG-USR-074 .
	SOC-ETH-TSD-021	Ensure cyclical review of the socio-ethical requirements throughout the lifecycle of AI and data manipulation functions and processes embedded in TRUSTEE. (this is pertinent especially where cumulative harm cannot be detected which does not exclude harmful impact, see "Veale & Borgesius (2021)	Desirable	EC Ethics Guidelines for Trustworthy AI;	LEG-TSD-015 LEG-TSD-017
	SOC-ETH-TSD-022	Provide opportunity for stakeholder participation and feedback collection to maintain TRUSTEE's trustworthiness. Such mechanisms should be easy to spot and access and ensure it elicits regular feedback from all stakeholders.	Desirable	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-013 , LEG-TSD-014 , LEG-TSD-110 .
	SOC-ETH-TSD-023	Provide and update codes of conduct that technical and non-technical stakeholders of Trustee such as developers and data providers are clear about and adhere to.	Desirable	EC Ethics Guidelines to Trustworthy AI	LEG-TSD-013 , LEG-TSD-014 , LEG-TSD-099 .
Ethical requirements through technical and non-technical	SOC-ETH-TSD-024	Ensure TRUSTEE system developers are clear about the regulatory obligations with regards to data manipulation and the deployment of AI functionalities	Mandatory	Refer to Legal Requirements for System Developers	LEG-TSD-029 ; LEG-USR-081-088 .

implementation	SOC-ETH-TSD-025	Ensure that TRUSTEE system developers adhere to clear design and data standards.	Mandatory	Refer to Legal Requirements for System Developers	<u>to some level LEG-TSD-086 and LEG-TSD-092 once AIA is in place including LEG-TSD-037 and LEG-USR-034 (users free to conduct DPIAs)</u>
	SOC-ETH-TSD-026	Ensure that design, data, and AI-related standards are up to date. ISO standards or IEEE standards series exist today, however future standards relating to AI and what it constitutes trustworthy AI and data manipulation may be developed. Such updates are necessary throughout the lifecycle of TRUSTEE and data shared on the platform to ensure that specific standards that the platform adheres to at the level of safety, technical robustness and transparency are met.	Desirable	IEEE Global Initiative on ethical designs of AI systems (https://standards.ieee.org/industry-connections/ec/autonomous-systems/)	<u>LEG-TSD-040</u>
	SOC-ETH-TSD-027	Provide labelling and certification that represent which standards have been met by the platform.	Optional	EC Ethics Guidelines for Trustworthy AI	<u>LEG-TSD-015</u> <u>LEG-TSD-019</u>

	SOC-ETH-TSD-028	Designate a person or team in charge of ethics issues relating to AI and data manipulation to provide proper governance around achieving ethical accountability. As part of this governance will be education, capacity building and awareness across all TRUSTEE stakeholders and developers to ensure ethical mind-set and ethical practices. As part of this governance and person/team responsible to oversee the ethical mind-set guides the development and growth of TRUSTEE will also be the role to ensure that there is diversity in the teams that design, update, and run the platform.	Desirable	AI Act	LEG-TSD-043 LEG-TSD-074 , LEG-TSD-083 .
	SOC-ETH-TSD-029	Report negative impacts in transparent and easy to understand language (respecting end users' conditions and abilities)	Optional	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-043 , LEG-TSD-074 , LEG-TSD-083 LEG-TSD-092 (relating to AIA)
	SOC-ETH-TSD-030	Provide and make explicit and easy to view mechanisms for redress especially with regards to vulnerable groups.	Optional	EU Agency for Fundamental Rights	LEG-TSD-095 , LEG-TSD-047
	SOC-ETH-TSD-031	Provide a clear 'whitelist' of rules, policies and procedures that TRUSTEE developers are guided by (and ensure transparency for end user's consumption) and provide a clear 'blacklist' restrictions and behaviours (AI and data manipulation that is prohibited or used in certain circumstances) which can be used for audit and compliance by outside third parties and for the maintenance of the platform's trustworthiness.	Desirable	AIA, Article 6(1) and (2); EC Ethics Guidelines for Trustworthy AI	LEG-TSD-103

	SOC-ETH-TSD-032	Provide explicability around levels of accuracy. That is, provide functionality around AI and data manipulation and their level of accuracy and the margins of error that any results might have.	Mandatory	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-035 , LEG-TSD-044 , LEG-TSD-048 .
	SOC-ETH-TSD-033	Provide transparency around the AI and data manipulation are functions with capabilities as well as limitations and they do not act as replacement of or on human individuals' behalf. That such functions are part of open data platforms such as TRUSTEE must be clearly communicated across to ensure that end users identify the AI functionalities as such.	Mandatory	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-002 , LEG-TSD-003 , LEG-TSD-015 , LEG-TSD-016 , LEG-TSD-040 , LEG-TSD-043 .
	SOC-ETH-TSD-034	Ensure societal and environmental wellbeing through explicability, and transparency around choice and alternative ways of achieving the same or similar results as the ones promised through data manipulation and AI use in open data platforms such as TRUSTEE	Mandatory	GDPR and EC Ethics Guidelines for Trustworthy AI	LEG-TSD-012 , LEG-TSD-044 , LEG-TSD-045 , LEG-TSD-047 , LEG-USR-080 .
Communication, documentation, and reporting for	SOC-ETH-TSD-035	Ensure protocols around access to data that is made available on open data platforms such as TRUSTEE are clear and put in place.	Mandatory	in line with GDPR	LEG-TSD-092

accountability and transparency Legal req for TRUSTEE system developers/providers: LEG-TSD-016 LEG-TSD-017 LEG-TSD-018; Addressing data consumer (see column C which raw relates to Data Consumer): LEG-USR-4, LEG-USR-6-10; LEG-USR-19-22, LEG-USR-46, LEG-USR-50-53, LEG-USR-55-56, LEG-USR-90	SOC-ETH-TSD-036	Provide clarity around the characteristics, limitations and potential shortcomings of the platform and its AI functionalities.	Desirable	EC Ethics Guidelines for Trustworthy AI	<u>LEG-TSD-003-007,</u> <u>LEG-TSD-012-014,</u> <u>LEG-TSD-043;</u> <u>LEG-USR-071-072.</u>
	SOC-ETH-TSD-037	Clearly communicate which audiences the platform and the data manipulation and AI functionalities are intended for.	Mandatory	EC Ethics Guidelines for Trustworthy AI	<u>LEG-TSD-013,</u> <u>LEG-TSD-107;</u> <u>LEG-USR-075.</u>
	SOC-ETH-TSD-038	Document and communicate benefits and opportunities from trustworthy open data platforms such as TRUSTEE with data manipulation and AI functionalities. As research advanced throughout the lifecycle of the project, this list will grow and a better understanding of the true potential of open data markets should be made available to stakeholders and various publics.	Mandatory	EC Ethics Guidelines for Trustworthy AI; https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682	<u>LEG-TSD-035,</u> <u>LEG-TSD-042,</u> <u>LEG-TSD-043,</u> <u>LEG-TSD-085.</u>
	SOC-ETH-TSD-039	Be able to explain the business model and exploitation of the platform and the rationale behind using AI functionalities.	Optional	EC Ethics Guidelines for Trustworthy AI	<u>LEG-TSD-092,</u> <u>LEG-TSD-093,</u> <u>LEG-TSD-095.</u>

	SOC-ETH-TSD-040	Define 'trustworthiness' to guide the socio-ethical assessment and evaluation of AI and data manipulation	Mandatory	Siau & Wang (2018)	LEG-TSD-013-014 , LEG-TSD-017 , LEG-TSD-050 , LEG-TSD-071 , LEG-TSD-098 , LEG-TSD-103 , LEG-TSD-109 .
	SOC-ETH-TSD-041	Ensure that prior to ethical and trustworthy assessment (i.e., the implementation of all of the above requirements), the TRUSTEE System Developers also meet the legally binding requirements (where relevant)	Mandatory	Refer to Legal Requirements for TRUSTEE System Developers	LEG-TSD-092-103
	SOC-ETH-TSD-042	Provide regular assessment of socio-ethical requirements throughout the lifecycle of the project to ensure that this list is updated, that education and awareness among key stakeholders (and especially amongst TRUSTEE developers) grows and improves, and that the mandatory and subsequently desirable requirements are met.	Desirable	EC Ethics Guidelines for Trustworthy AI; more broadly also refer to CNIL's 6th principle: "supervising continuous improvement"	LEG-TSD-111-113 , and to AIA LEG-TSD-092-093

6.2 SOCIO-ETHICAL REQUIREMENTS FOR TRUSTEE END-USERS¹⁰³⁸

Nature of requirement	Req No	Personas (Data Provider, Model Provider, Consumer, Developer)	Requirement (ethical principles and ethical imperatives) (Technical, Management, Legal teams + AI Model Providers)	Priority (Must-have. Mandatory Requirement, Should-have. Desirable Requirement, Could-have. Optional Requirement, Will-not-have. Possible Future Enhancement)	Legal Act/ Guideline/Standard Relevant	No of related legal requirements for TRUSTEE system developers and end users

¹⁰³⁸ *points relating to mapping (from about LEG-USR-055 to about LEG-USR-068, LEG-USR-077 and LEG-USR-078)
any aspects relating to legal req for end user and re-use of data
in relation to TRUSTEE use by public sector body
in relation to intellectual property rights (LEG-TSD-112-113 and in LEG-USR-061, 094, 096 and 097)

Adhering to fundamental human rights	SOC-ETH-USR-001	Data consumer should be clearly informed about the fundamental human rights of those whose data is shared on platforms such as TRUSTEE and those who may be impacted by the data processing. Data consumers should be clearly informed how the platform meets these rights.	Data consumer should be made aware and clearly informed that any data or data manipulation and AI use should respect human dignity (i.e., treat humans as 'moral subjects' not as 'objects' that can be scored, manipulated etc.	Mandatory	EU Treaties, EU Charter and International fundamental human rights (Article 1); EC expert guidelines for trustworthy AI; HRESIA Model (Mantelero, 2022);	LEG-TSD-002, LEG-TSD-003, LEG-TSD-013, LEG-TSD-014, LEG-TSD-040, LEG-TSD-043, LEG-TSD-047; LEG-USR-002, LEG-USR-025, LEG-USR-036-037 (in relation to fundamental rights), LEG-USR-059, LEG-USR-062, LEG-USR-092
	SOC-ETH-USR-002		Data consumer should be informed of the fundamental human rights of data subjects of data being processing ed via TRUSTEE services and avoid making automated decisions which may have legal effect or similarly affect the data subjects behind encrypted data, and TRUSTEE must be able to provide clear choices and alternatives to the data manipulation and AI functionalities offered.	Mandatory	in relation to GDPR, Article 22(1 2)	
	SOC-ETH-USR-003		Data consumer should be made aware and clearly informed that the system (e.g., TRUSTEE) is designed to respect democracy, justice, and the rule of law, plurality of choice, values, and decision-making of individuals.	Mandatory	EC Ethics Guidelines for Trustworthy AI	
	SOC-ETH-USR-004		Data consumer should be made aware that the platform respects equality, non-discrimination, and solidarity whereby datasets used to train AI models strive to consider diversity and be inclusive of diverse populations.	Mandatory	EC Ethics Guidelines for Trustworthy AI	

	SOC-ETH-USR-005		Data consumer should be made aware the platform is designed to respect citizens' rights with regards to deploying AI onto TRUSTEE	Mandatory	EU and fundamental human rights; EC expert guidelines for trustworthy AI; HRESIA Model; DPIA; ICO Ethics framework (2020) https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020	
	SOC-ETH-USR-006		Data consumer should be made aware that the platform is designed to ensure that any AI and data manipulation deployed is done for the purpose* of enhancing and benefiting human and societal wellbeing and human capabilities.	Mandatory	EC expert guidelines for trustworthy AI; AIA; *Any AI deployment should have a clearly defined purpose - principle 1, CNIL (2022)	
	SOC-ETH-USR-007		Data consumer should be made aware that the platform aims to provide safeguards against overconfidence and overreliance on any of the data manipulation AI provisions on TRUSTEE.	Desirable	EC Ethics Guidelines for Trustworthy AI	
	SOC-ETH-USR-008		Data consumer should be made aware that the platform aims to prioritise sustainable development, personal data rights, and agency over digital identity; and ensure legal frameworks for accountability are met.	Desirable	EC Ethics Guidelines for Trustworthy AI	

	SOC-ETH-USR-009	ALL	Any data or data manipulation and AI use should respect human dignity (i.e., treat humans as 'moral subjects' not as 'objects' that can be scored, manipulated etc.	Mandatory	EU Treaties, EU Charter and International fundamental human rights (Article 1); EC expert guidelines for trustworthy AI; HRESIA Model (Mantelero, 2022);	
Communication, documentation, and reporting for accountability and transparency Legal req for TRUSTEE system developers/providers : LEG-TSD-016 LEG-TSD-017 LEG-TSD-018; Addressing data consumer (see column C which raw relates to Data Consumer): LEG-USR-4, LEG-USR-6-10; LEG-USR-19-22, LEG-USR-46, LEG-USR-50-53,	SOC-ETH-USR-010	Developer, Model provider	Report negative impacts in transparent and easy to understand language (respecting end users' conditions and abilities)	Optional	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-035 , LEG-TSD-044 , LEG-TSD-048 .
	SOC-ETH-USR-011	Data consumer	Provide and make explicit and easy to view mechanisms for redress especially with regards to vulnerable groups.	Optional	EU Agency for Fundamental Rights	LEG-TSD-002 , LEG-TSD-003 , LEG-TSD-015 , LEG-TSD-016 , LEG-TSD-040 , LEG-TSD-043 .
	SOC-ETH-USR-012	Consumer (when using the obtained results)	Provide explicability around levels of accuracy. That is, provide functionality around AI and data manipulation and their level of accuracy and the margins of error that any results might have.	Mandatory	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-092

LEG-USR-55-56, LEG-USR-90	SOC-ETH-USR-013	Consumer (when using the obtained results)	Provide transparency around the AI and data manipulation are functions with capabilities as well as limitations and they do not act as replacement of or on human individuals' behalf. That such functions are part of open data platforms such as TRUSTEE must be clearly communicated across to ensure that end users identify the AI functionalities as such.	Mandatory	EC Ethics Guidelines for Trustworthy AI	LEG-TSD-003-007 , LEG-TSD-012-014 , LEG-TSD-043 , LEG-USR-071-072 .
	SOC-ETH-USR-014	Data consumer (in terms of finding alternatives and responsible use)	Ensure societal and environmental wellbeing through explicability, and transparency around choice and alternative ways of achieving the same or similar results as the ones promised through data manipulation and AI use in open data platforms such as TRUSTEE	Mandatory	GDPR and EC Ethics Guidelines for Trustworthy AI	LEG-TSD-013 , LEG-TSD-107 , LEG-USR-075 .
	SOC-ETH-USR-015	Data consumer (when using the obtained results)	Ensure protocols around access to data that is made available on open data platforms such as TRUSTEE are clear and put in place.	Mandatory	in line with GDPR	LEG-TSD-035 , LEG-TSD-042 , LEG-TSD-043 , LEG-TSD-085 .
	SOC-ETH-USR-016	All	Document and communicate benefits and opportunities from trustworthy open data platforms such as TRUSTEE with data manipulation and AI functionalities. As research advanced throughout the lifecycle of the project, this list will grow and a better understanding of the true potential of platforms such as TRUSTEE should be made available to stakeholders and various publics.		EC Ethics Guidelines for Trustworthy AI; https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682	LEG-TSD-092-103

