



D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment

Document Identification			
Status	Final	Due Date	31/08/2023
Version	2.0	Submission Date	07/06/2024

Related WP	WP2	Document Reference	D2.3
Related Deliverable(s)	D2.1, D2.2	Dissemination Level (*)	PU
Lead Participant	INTRA	Lead Author	Panos Matzakos, INTRA
Contributors	UTH, ATOS, DBC, FHG, NOR, VTT, FUJ_DE, FUJ_LU, UMU, UOM, DUT, UoG, CEA, SQD, XLAB, QBE, LSTECH, EXUS	Reviewers	Matija Cankar, XLAB
			Antonis Chronakis, SVI

Keywords:

System requirements, functional & non-functional requirements, security requirements, technical specifications, platform architecture, social & legal assessment, Architecture, Data Usage Control, Data Sovereignty, Dataspaces.

Disclaimer for Deliverables with dissemination level PUBLIC

This document is issued within the frame and for the purpose of the TANGO project. This project has received funding from the European Union's Horizon Europe Framework Programme under Grant Agreement No. 101070052. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the European Commission.

The dissemination of this document reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains. **This deliverable is subject to final acceptance by the European Commission.**

This document and its content are the property of the TANGO Consortium. The content of all or parts of this document can be used and distributed provided that the TANGO project and the document are properly referenced.

Each TANGO Partner may use this document in conformity with the TANGO Consortium Grant Agreement provisions.

(*) Dissemination level: **(PU)** Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page). **(SEN)** Sensitive, limited under the conditions of the Grant Agreement. **(Classified EU-R)** EU RESTRICTED under the Commission Decision No2015/444. **(Classified EU-C)** EU CONFIDENTIAL under the Commission Decision No2015/444. **(Classified EU-S)** EU SECRET under the Commission Decision No2015/444.

Document Information

List of Contributors	
Name	Partner
Panos Matzakos	INTRA
Maria Fritzela	INTRA
Eleni Veroni	INTRA
Apostolos Apostolaras	UTH
Ilias Syrigos	UTH
Stavroula Maglavera	UTH
Tomás Pariente Lobo	ATOS
Ross Little Armitt	ATOS
Sergio Salmerón	ATOS
Nikos Avgerinos	DBC
Sotirios Michagiannis	DBC
Vasiliki Andriopoulou	DBC
Kristina Thim	NOR
Vitalii Demianets	NOR
Felix Hermsen	FHG
Ville Ollikainen	VTT
Erik Hieta-aho	VTT
Anni Karinsalo	VTT
Sami Lehtonen	VTT
Valtteri Lipiäinen	VTT
Jürgen Neises	FUJ_DE
Sofiane Lagraa	FUJ_LU
Moussa Ouedraogo	FUJ_LU
Sven Rasmusen	FUJ_LU
Jesus Garcia	UMU
Antonio Skarmeta	UMU
Maria Hernandez Padilla	UMU
Athanasios Stratikopoulos	UOM
Kaitai Liang	DUT
Sakshyam Panda	UoG
Emmanouil Panaousis	UoG
Nicolas Belleville	CEA
John Zaras	SQD
Jakob Jenko	XLAB
Niklas Palaghias	QBE
George Sachpatzidis	QBE
Evangelos Kotsifakos	LSTECH
Konstantinos Kentrotis	EXUS
Cristina Nichiforov	EXUS

Document History			
Version	Date	Change editors	Changes
0.01	07/03/2023	INTRA	Initial ToC.
0.1	27/03/2023	INTRA	Updated ToC, incorporating input from UTH & DBC.
0.2	29/05/2023	INTRA	Merged initial inputs for all TANGO components, Section 8 restructure.
0.3	13/06/2023	INTRA	Merged updates for all TANGO components functionality descriptions.
0.4	26/07/2023	INTRA	Merged FR/NFR requirements for all TANGO components, inputs on the architecture and TANGO integrated platform, as well as on Privacy, Ethical, Social and Legal Impact Assessment.
0.5	19/09/2023	INTRA	Merged inputs on Abstract, introduction, concept and methodology, as well as updated inputs in TANGO components sections, architecture and privacy, ethical and legal impact assessment.
0.6	28/09/2023	INTRA	Merged updates relative to the involvement of some technical components to the pilots and inputs to conclusions; Updated content based on the comments from the internal reviews.
0.7	29/09/2023	INTRA	Released deliverable version ready for final quality check.
1.0	03/10/2023	ATOS / FUJ_DE	Final version for submission
2.0	07/06/2024	INTRA	Revised version with updates in figures 1, 3, 4, 5, 9, 11, 12, 13, 14, 22, 24, 25, 31, 44, 45 to improve their quality, according to the review report.

Quality Control		
Role	Who	Approval Date
Deliverable leader	Panos Matzakos (INTRA)	29/09/2023
Quality manager	Jürgen Neises (FSDE)	02/10/2023
Project Coordinator	Tomás Pariente Lobo (ATOS)	03/10/2023

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	3 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table of Contents

Document Information	2
Table of Contents	4
List of Tables.....	9
List of Figures	14
List of Acronyms.....	15
Executive Summary	17
1 Introduction	18
1.1 Purpose of the document.....	18
1.2 Relation to other project work.....	18
1.3 Structure of the document	19
2 Concept & Methodology	20
2.1 TANGO Concept & Key Technological Areas.....	20
2.1.1 Distributed Privacy-preserving Data Management and Storage	20
2.1.2 Distributed Trust Management Framework	21
2.1.3 AI-based Framework for Green & Trustworthy Operations	23
2.1.4 TANGO Integrated Platform Concepts	24
2.2 Methodological Framework	25
2.2.1 Project Methodology	25
2.2.2 Requirements Methodology	26
2.2.3 Architecture Methodology.....	29
3 Distributed Privacy-preserving Data Management and Storage.....	30
3.1 Blockchain-based Data Storage and Sharing [T3.1]	30
3.1.1 Description & Functionality	30
3.1.2 Associated Pilot Use Cases	34
3.1.3 Threat Landscape.....	35
3.1.4 Functional & Non-functional Requirements	36
3.2 Trustworthy Data Sharing [T3.2]	37
3.2.1 Description & Functionality	37
3.2.2 Associated Pilot Use Cases	43
3.2.3 Threat Landscape.....	44
3.2.4 Functional & Non-functional Requirements	44
3.3 Confidentiality and Privacy by Design [T3.3]	45
3.3.1 Description & Functionality	45
3.3.2 Associated Pilot Use Cases	49
3.3.3 Threat Landscape.....	50

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	4 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

3.3.4	Functional & Non-functional Requirements	50
3.4	Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms [T3.4]	51
3.4.1	Description & Functionality	51
3.4.2	Associated Pilot Use Cases	54
3.4.3	Threat Landscape.....	54
3.4.4	Functional & Non-functional Requirements	55
4	Distributed Trust Management Framework.....	56
4.1	Self-sovereign Identity Management [T4.1]	56
4.1.1	Description & Functionality	56
4.1.2	Associated Pilot Use Cases	59
4.1.3	Threat Landscape.....	60
4.1.4	Functional & Non-functional Requirements	61
4.2	Seamless Onboarding for Users and Devices [T4.2]	62
4.2.1	Description & Functionality	62
4.2.2	Associated Pilot Use Cases	64
4.2.3	Threat Landscape.....	65
4.2.4	Functional & Non-functional Requirements	65
4.3	User continuous behavioural authentication [T4.3]	66
4.3.1	Description & Functionality	66
4.3.2	Associated Pilot Use Cases	67
4.3.3	Threat Landscape.....	68
4.3.4	Functional & Non-functional Requirements	68
4.4	Device continuous behavioural authentication [T4.4]	69
4.4.1	Description & Functionality	69
4.4.2	Associated Pilot Use Cases	72
4.4.3	Threat Landscape.....	73
4.4.4	Functional & Non-functional Requirements	73
4.5	Hardening Against Side-channel Attacks [T4.5]	74
4.5.1	Description & Functionality	74
4.5.2	Associated Pilot Use Cases	76
4.5.3	Threat Landscape.....	76
4.5.4	Functional & Non-functional Requirements	76
5	AI-based Framework for Green & Trustworthy Operations	78
5.1	Exploratory Data Analysis Engine [T5.1].....	78
5.1.1	Description & Functionality	78
5.1.2	Associated Pilot Use Cases	83

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	5 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

5.1.3	Threat Landscape.....	84
5.1.4	Functional & Non-functional Requirements	84
5.2	Energy efficient AI model training [T5.2]	85
5.2.1	Description & Functionality	85
5.2.2	Associated Pilot Use Cases	91
5.2.3	Threat Landscape.....	92
5.2.4	Functional & Non-functional Requirements	93
5.3	Dynamic Intelligent Execution on Heterogeneous Systems [T5.3]	94
5.3.1	Description & Functionality	94
5.3.2	Associated Pilot Use Cases	96
5.3.3	Threat Landscape.....	97
5.3.4	Functional & Non-functional Requirements	97
5.4	Privacy Threat Modelling & Identification for Trustworthy AI [T5.4]	97
5.4.1	Description & Functionality	97
5.4.2	Associated Pilot Use Cases	100
5.4.3	Threat Landscape.....	101
5.4.4	Functional & Non-functional Requirements	101
5.5	X-AI For Privacy and Trust Enhancement [T5.5].....	102
5.5.1	Description & Functionality	102
5.5.2	Significance and Added Value for TANGO.....	104
5.5.3	Associated Pilot Use Cases	104
5.5.4	Threat Landscape.....	105
5.5.5	Functional & Non-functional Requirements	105
5.6	Infrastructure Management based on AI [T5.6].....	105
5.6.1	Description & Functionality	105
5.6.2	Associated Pilot Use Cases	108
5.6.3	Threat Landscape.....	109
5.6.4	Functional & Non-functional Requirements	109
6	Visualisation Techniques & User Interfaces	110
6.1	User Interfaces [T6.3]	110
6.1.1	Description & Functionality	110
6.1.2	Significance and Added Value for TANGO.....	112
6.1.3	Associated Pilot Use Cases	112
6.1.4	Threat Landscape.....	112
6.1.5	Functional & Non-functional Requirements	112
6.2	Visualisation Techniques [T6.4]	113
6.2.1	Description & Functionality	113

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	6 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

6.2.2	Associated Pilot Use Cases	114
6.2.3	Functional & Non-functional Requirements	115
7	TANGO Reference Architecture	117
7.1	Introduction	117
7.1.1	TANGO Architecture Overview	117
7.1.2	GAIA-X and IDSA compatibility.....	118
7.1.3	Enabling Pilots with the TANGO Core Architecture	119
7.2	TANGO Connector	119
7.2.1	Connection Manager	120
7.2.2	Configuration Manager	121
7.2.3	Audit Log.....	122
7.2.4	Usage Control.....	122
7.2.5	Message Router	125
7.3	Blueprint.....	126
7.4	TANGO Management Connector	128
7.4.1	Connection Manager	129
7.4.2	Deployment Orchestrator	129
7.4.3	Blueprint Repository	130
7.4.4	Clearing House	130
7.4.5	Verifiable Data Registry.....	131
7.4.6	Decentralized Identifiers support without a VDR	131
7.4.7	Decentralized identifiers supported with a VDR (Fides-based approach)	132
7.4.8	Trustworthiness Module.....	133
7.4.9	User Behaviour Exchange Module.....	133
7.5	TANGO App Store	133
7.5.1	Portal	134
7.5.2	Image Repository.....	134
7.6	Example Workflow	134
7.6.1	Steps	139
8	TANGO Integrated Platform [T6.1-2].....	143
8.1	Description of production environment	143
8.2	Non-functional Requirements	145
9	Privacy, Ethical, Social and Legal Impact Assessment	147
9.1	Definitions.....	147
9.2	Introduction.....	147
9.3	Data mapping process	148

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	7 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

9.4 Description of the data processing activities during the development of the TANGO framework.....	150
9.5 Lawfulness of the data processing	156
9.5.1 Key provisions of GDPR.....	156
9.5.2 Implementation of those provisions to the data processing for TANGO	158
9.6 Implemented measures related to the data subjects.....	159
9.6.1 Key provisions of GDPR.....	160
9.6.2 Currently implemented measures	160
9.6.3 Actions in case of a data breach event.....	161
9.7 Key risks	161
9.8 Technical and operational security measures.....	162
9.8.1 Key provisions of GDPR.....	162
9.8.2 Currently implemented security measures	162
9.9 Assessment related to the architecture of the TANGO framework.....	167
9.10 Ethics assessment.....	169
9.11 Social impact assessment	171
9.12 Considerations.....	172
9.13 Key legal requirements in terms of privacy and ethics	173
10 Conclusions	174
11 Bibliography.....	175
Annexes	179
A. System Requirements for Distributed Privacy-preserving Data Management and Storage	179
A.1 Functional Requirements.....	179
A.2 Non-Functional Requirements	186
B. System Requirements for Distributed Trust Management Framework.....	191
B.1 Functional Requirements.....	191
B.2 Non-Functional Requirements.....	200
C. System Requirements for AI-based Framework for Green & Trustworthy Operations.....	208
C.1 Functional Requirements.....	208
C.2 Non-Functional Requirements.....	219
D. System Requirements for Visualisation Techniques & User Interfaces	225
D.1 Functional Requirements.....	225
D.2 Non-Functional Requirements	229
E. System Requirements for the TANGO Integrated Platform	232
E.1 Non-Functional Requirements.....	232
F. Privacy, Ethics, Legal, and Societal Requirements	236

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	8 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

List of Tables

Table 1: Template for each Functional Requirement description	27
Table 2: Defined Categories of Non-Functional Requirements	28
Table 3: Association of Pilot Use Cases with Fides component T3.1	34
Table 4: Blockchain-based Data Storage and Sharing Component Functional Requirements (Author: NOR)	36
Table 5: Blockchain-based Data Storage and Sharing Component Non-Functional Requirements (Author: NOR)	36
Table 6: Association of Pilot Use Cases with trustworthiness scoring module of T3.2	44
Table 7: Association of Pilot Use Cases with UPCV module of T3.2	44
Table 8: Trustworthy Data Sharing Component Functional Requirements (Author: FHG, FUJ_DE, VTT)	44
Table 9: Trustworthy Data Sharing Component Non-Functional Requirements (Author: FHG, FUJ_DE, VTT)	45
Table 10: Association of Pilot Use Cases with confidentiality and privacy by design functionalities of T3.3	49
Table 11: Confidentiality and Privacy by design Component Functional Requirements (Author: UMU)	51
Table 12: Confidentiality and Privacy by Design Non-Functional Requirements (Author: UMU)	51
Table 13: Association of Pilot Use Cases with self-encryption and decryption techniques of T3.4	54
Table 14: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms Functional Requirements (Author: DUT)	55
Table 15: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms Non-Functional Requirements (Author: DUT)	55
Table 16: Association of Pilot Use Cases with SSI functionalities of T4.1	60
Table 17: Self-sovereign identity management Functional Requirements (Author: VTT, UMU, ATOS)	61
Table 18 Self-sovereign identity management Non-Functional Requirements (Author: VTT, UMU, ATOS)	61
Table 19: Association of Pilot Use Cases with SSI Onboard component of T4.2	64
Table 20: Seamless Onboarding for Users and Devices Functional Requirements (Author: QBE)	65
Table 21: Seamless Onboarding for Users and Devices Non-Functional Requirements (Author: QBE)	66
Table 22: Association of Pilot Use Cases with UCBA component of T4.3	68
Table 23: User continuous behavioural authentication Functional Requirements (Author: QBE)	68
Table 24 User continuous behavioural authentication Non-Functional Requirements (Author: QBE)	69
Table 25: Association of Pilot Use Cases with DCBA component of T4.4	72
Table 26: Device continuous behavioural authentication Functional Requirements (Author: UTH)	73
Table 27: Device continuous behavioural authentication Non-Functional Requirements (Author: UTH)	74
Table 28: Association of Pilot Use Cases with LLVM-based compiler of T4.5	76
Table 29: Hardening Against Side-channel Attacks Functional Requirements (Author: CEA)	76
Table 30: Hardening Against Side-channel Attacks Non-Functional Requirements (Author: CEA)	76
Table 31: Association of Pilot Use Cases with T5.1	83
Table 32: EDAE Component Functional Requirements (Author: EXUS)	84
Table 33: EDAE Component Non-Functional Requirements (Author: EXUS)	85
Table 34: Association of Pilot Use Cases with the Federated Learning Component of T5.2	91
Table 35: Association of Pilot Use Cases with the MLOps Component of T5.2	92
Table 36: Association of Pilot Use Cases with the AutoML Component of T5.2	92
Table 37: Federated Learning Component Functional Requirements (Author: ATOS)	93
Table 38: MLOps Component Functional Requirements (Author: ATOS)	93
Table 39: AutoML Component Functional Requirements (Author: ATOS)	93
Table 40: Federated Learning Component Non-Functional Requirements (Author: ATOS)	94
Table 41: MLOps Component Non-Functional Requirements (Author: ATOS)	94
Table 42: AutoML Component Non-Functional Requirements (Author: ATOS)	94

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	9 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Table 43: Association of Pilot Use Cases with the TornadoVM Component of T5.3	96
Table 44: TornadoVM Component Functional Requirements (Author: UOM)	97
Table 45: TornadoVM Non-Functional Requirements (Author: UOM)	97
Table 46: Association of Pilot Use Cases with the PEC Component of T5.4	100
Table 47: Association of Pilot Use Cases with the PAT Component of T5.4	100
Table 48: Privacy Enhancing Component Functional Requirements (Author: UOG)	101
Table 49: Privacy Assurance Tool Functional Requirements (Author: FUJ_LU)	101
Table 50: Privacy Enhancing Component Non-Functional Requirements (Author: UOG)	102
Table 51: Privacy Assurance Tool Non-Functional Requirements (Author: UOG)	102
Table 52: Association of Pilot Use Cases with the XAI Component of T5.5	104
Table 53: XAI Component Functional Requirements (Author: SQD)	105
Table 54: XAI Component Non-Functional Requirements (Author: SQD)	105
Table 55: Categorization of Energy-Intensive Processing Tasks	107
Table 56: Association of Pilot Use Cases with the RENOPS Component of T5.6	108
Table 57: RENOPS Component Functional Requirements (Author: XLAB)	109
Table 58: RENOPS Component Non-Functional Requirements (Author: XLAB)	109
Table 59: User Interfaces Functional Requirements (Author: LST)	112
Table 60: User Interfaces Non-Functional Requirements (Author: LST)	113
Table 61: Association of Pilot Use Cases with the Visualisation and Reporting Techniques Component of T6.4	115
Table 62: Visualization and Reporting Techniques Component Functional Requirements (Author: SQD)	115
Table 63: Visualization and Reporting Techniques Component Non-Functional Requirements (Author: SQD)	116
Table 64: IDSA and GXFS alternative implementations	119
Table 65: TANGO Integrated platform Non-Functional Requirements (Author: INTRA, NOR)	145
Table 66: Personal data processed per TANGO partner	151
Table 67: Privacy, Ethics, Legal, and Societal Requirements (Author: DBC)	173
Table 68: Blockchain-based Data Storage and Sharing: FR-DATA-BDSS-001	179
Table 69: Blockchain-based Data Storage and Sharing: FR-DATA-BDSS-002	179
Table 70: Blockchain-based Data Storage and Sharing: FR-DATA-BDSS-003	179
Table 71: Trustworthy Data Sharing: FR-DATA-TDS-001	180
Table 72: Trustworthy Data Sharing: FR-DATA-TDS-002	180
Table 73: Trustworthy Data Sharing: FR-DATA-TDS-003	181
Table 74: Trustworthy Data Sharing: FR-DATA-TDS-004	181
Table 75: Trustworthy Data Sharing: FR-DATA-TDS-005	182
Table 76: Confidentiality and Privacy by Design: FR-DATA-CPD-001	182
Table 77: Confidentiality and Privacy by Design: FR-DATA-CPD-002	182
Table 78: Confidentiality and Privacy by Design: FR-DATA-CPD-003	183
Table 79: Confidentiality and Privacy by Design: FR-DATA-CPD-004	183
Table 80: Confidentiality and Privacy by Design: FR-DATA-CPD-005	184
Table 81: Confidentiality and Privacy by Design: FR-DATA-CPD-006	184
Table 82: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: FR-DATA-SED-T-001	184
Table 83: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: FR-DATA-SED-T-002	185
Table 84: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: FR-DATA-SED-T-003	185
Table 85: Blockchain-based Data Storage and Sharing: NFR-PER-BDSS-001	186
Table 86: Blockchain-based Data Storage and Sharing: NFR-REL-BDSS-001	186
Table 87: Blockchain-based Data Storage and Sharing: NFR-SEC-BDSS-001	186
Table 88: Trustworthy Data Sharing: NFR-SEC-TDS-001	187
Table 89: Trustworthy Data Sharing: NFR-SEC-TDS-002	187
Table 90: Trustworthy Data Sharing: NFR-SEC-TDS-003	187
Table 91: Trustworthy Data Sharing: NFR-UI-TDS-001	187
Table 92: Trustworthy Data Sharing: NFR-UI-TDS-002	188
Table 93: Confidentiality and Privacy by Design: NFR-SEC-CPD-001	188

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	10 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Table 94: Confidentiality and Privacy by Design: NFR-SEC-CPD-002	188
Table 95: Confidentiality and Privacy by Design: NFR-PER-CPD-001	189
Table 96: Confidentiality and Privacy by Design: NFR-REL-CPD-001	189
Table 97: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: NFR-MAIN-SED-001	189
Table 98: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: NFR-PER-SED-001	190
Table 99: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: NFR-COMP-SED-001	190
Table 100: Self-sovereign identity management: FR-TRUST-SSIM-001	191
Table 101: Self-sovereign identity management: FR-TRUST-SSIM-002	191
Table 102: Self-sovereign identity management: FR-TRUST-SSIM-003	192
Table 103: Self-sovereign identity management: FR-TRUST-SSIM-004	192
Table 104: Self-sovereign identity management: FR-TRUST-SSIM-005	192
Table 105: Self-sovereign identity management: FR-TRUST-SSIM-006	193
Table 106: Self-sovereign identity management: FR-TRUST-SSIM-007	193
Table 107: Self-sovereign identity management: FR-TRUST-SSIM-008	194
Table 108: Self-sovereign identity management: FR-TRUST-SSIM-009	194
Table 109: Self-sovereign identity management: FR-TRUST-SSIM-010	195
Table 110: Self-sovereign identity management: FR-TRUST-SSIM-011	195
Table 111: Self-sovereign identity management: FR-TRUST-SSIM-012	195
Table 112: Seamless Onboarding for Users and Devices: FR-TRUST-SOUD-001	196
Table 113: Seamless Onboarding for Users and Devices: FR-TRUST-SOUD-002	196
Table 114: User continuous behavioural authentication: FR-TRUST-UCBA-001	196
Table 115: User continuous behavioural authentication: FR-TRUST-UCBA-002	197
Table 116: User continuous behavioural authentication: FR-TRUST-UCBA-003	197
Table 117: User continuous behavioural authentication: FR-TRUST-UCBA-004	197
Table 118: User continuous behavioural authentication: FR-TRUST-UCBA-005	198
Table 119: Device continuous behavioural authentication: FR-TRUST-DCBA-001	198
Table 120: Device continuous behavioural authentication: FR-TRUST-DCBA-002	198
Table 121: Device continuous behavioural authentication: FR-TRUST-DCBA-003	198
Table 122: Device continuous behavioural authentication: FR-TRUST-DCBA-004	199
Table 123: Device continuous behavioural authentication: FR-TRUST-DCBA-005	199
Table 124: Device continuous behavioural authentication: FR-TRUST-DCBA-006	199
Table 125: Hardening against side-channel attacks: FR-TRUST-HASA-001	200
Table 126: Self-sovereign identity management: NFR-SEC-SSIM-001	200
Table 127: Self-sovereign identity management: NFR-SEC-SSIM-002	201
Table 128: Self-sovereign identity management: NFR-SEC-SSIM-003	201
Table 129: Self-sovereign identity management: NFR-SEC-SSIM-004	201
Table 130: Self-sovereign identity management: NFR-PER-SSIM-001	202
Table 131: Self-sovereign identity management: NFR-REL-SSIM-001	202
Table 132: Self-sovereign identity management: NFR-REL-SSIM-002	202
Table 133: Self-sovereign identity management: NFR-COMP-SSIM-001	203
Table 134: Self-sovereign identity management: NFR-SEC-SSIM-005	203
Table 135: Seamless Onboarding for Users and Devices: NFR-COMP-SOUD-001	203
Table 136: Seamless Onboarding for Users and Devices: NFR-PER-SOUD-001	203
Table 137: Seamless Onboarding for Users and Devices: NFR-EXIST-SOUD-001	204
Table 138: User continuous behavioural authentication: NFR-COMP-UCBA-001	204
Table 139: User continuous behavioural authentication: NFR-REL-UCBA-001	204
Table 140: User continuous behavioural authentication: NFR-EXIST-UCBA-001	204
Table 141: User continuous behavioural authentication: NFR-REL-UCBA-002	205
Table 142: Device continuous behavioural authentication: NFR-COMP-DCBA-001	205
Table 143: Device continuous behavioural authentication: NFR-PER-DCBA-001	205
Table 144: Device continuous behavioural authentication: NFR-REL-DCBA-001	206
Table 145: Device continuous behavioural authentication: NFR-SEC-DCBA-001	206
Table 146: Hardening against side-channel attacks: NFR-EXIST-HASA-001	206
Table 147: Hardening against side-channel attacks: NFR-COMP-HASA-001	207

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	11 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Table 148: Hardening against side-channel attacks: NFR-SEC-HASA-001	207
Table 149: Exploratory Data Analysis Engine: FR-AI-EDAE-001	208
Table 150: Exploratory Data Analysis Engine: FR-AI-EDAE-002	208
Table 151: Exploratory Data Analysis Engine: FR-AI-EDAE-003	208
Table 152: Exploratory Data Analysis Engine: FR-AI-EDAE-004	209
Table 153: Exploratory Data Analysis Engine: FR-AI-EDAE-005	209
Table 154: Exploratory Data Analysis Engine: FR-AI-EDAE-006	209
Table 155: Exploratory Data Analysis Engine: FR-AI-EDAE-007	209
Table 156: Exploratory Data Analysis Engine: FR-AI-EDAE-008	210
Table 157: Exploratory Data Analysis Engine: FR-AI-EDAE-009	210
Table 158: Exploratory Data Analysis Engine: FR-AI-EDAE-010	210
Table 159: Exploratory Data Analysis Engine: FR-AI-EDAE-011	210
Table 160: Exploratory Data Analysis Engine: FR-AI-EDAE-012	211
Table 161: Exploratory Data Analysis Engine: FR-AI-EDAE-013	211
Table 162: Federated Learning Component: FR-AI-FL-001	211
Table 163: Federated Learning Component: FR-AI-FL-002	212
Table 164: Federated Learning Component: FR-AI-FL-003	212
Table 165: MLOps Component: FR-AI-MLOPS-001	212
Table 166: MLOps Component: FR-AI-MLOPS-002	213
Table 167: AutoML Component: FR-AI-AUTOML-001	213
Table 168: AutoML Component: FR-AI-AUTOML-002	213
Table 169: TornadoVM: FR-AI-TVM-001	214
Table 170: TornadoVM, FR-AI-TVM-002	214
Table 171: Privacy Enhancing Component: FR-AI-PEC-001	214
Table 172: Privacy Enhancing Component: FR-AI-PEC-002	215
Table 173: Privacy Enhancing Component: FR-AI-PEC-003	215
Table 174: Privacy Assurance Tool: FR-AI-PAT-001	215
Table 175: Privacy Assurance Tool: FR-AI-PAT-002	215
Table 176: Privacy Assurance Tool: FR-AI-PAT-003	216
Table 177: Privacy Assurance Tool: FR-AI-PAT-004	216
Table 178: Privacy Assurance Tool: FR-AI-PAT-005	216
Table 179: Privacy Assurance Tool: FR-AI-PAT-006	217
Table 180: Privacy Assurance Tool: FR-AI-PAT-007	217
Table 181: XAI Component: FR-AI-XAI-001	217
Table 182: XAI Component: FR-AI-XAI-002	217
Table 183: RENOPS Component: FR-AI-RENOPS-001	218
Table 184: RENOPS Component: FR-AI-RENOPS-002	218
Table 185: RENOPS Component: FR-AI-RENOPS-003	218
Table 186: RENOPS Component: FR-AI-RENOPS-004	218
Table 187: Exploratory Data Analysis Engine: NFR-COMP -EDAE-001	219
Table 188: Exploratory Data Analysis Engine: NFR-PORT-EDAE-001	219
Table 189: Federated Learning Component: NFR-PER-FL-001	219
Table 190: Federated Learning Component: NFR-PER-FL-002	220
Table 191: Federated Learning Component: NFR-SEC-FL-001	220
Table 192: MLOps Component: NFR-PORT-MLOPS-001	220
Table 193: MLOps Component: NFR-UI-MLOPS-001	220
Table 194: AutoML Component, NFR-PER-AUTOML-001	221
Table 195: TornadoVM: NFR-COMP-TVM-001	221
Table 196: Privacy Enhancing Component: NFR-COMP-PEC-001	222
Table 197 Privacy Enhancing Component: NFR-UI-PEC-001	222
Table 198: Privacy Assurance Tool: NFR-SEC-PAT-001	222
Table 199: Privacy Assurance Tool: NFR-UI-PAT-001	222
Table 200: XAI Component: NFR-COMP-XAI-001	223
Table 201: XAI Component: NFR-PER-XAI-001	223
Table 202: XAI Component: NFR-EXIST-XAI-002	223
Table 203: XAI Component: NFR-PORT-XAI-001	223
Table 204: XAI Component: NFR-REL-XAI-001	224

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	12 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 205: RENOPS Component: NFR-MAIN-RENOPS-001	224
Table 206: RENOPS Component: NFR-PORT-RENOPS-001	224
Table 207: RENOPS Component: NFR-REL-RENOPS-001	224
Table 208: RENOPS Component: NFR-SEC-RENOPS-001	225
Table 209: User Interfaces: FR-S-UI-001	225
Table 210: User Interfaces: FR-S-UI-002	225
Table 211: User Interfaces: FR-S-UI-003	226
Table 212: User Interfaces: FR-S-UI-004	226
Table 213: Visualisation and Reporting Techniques Component: FR-S-VIS-001	226
Table 214: Visualisation and Reporting Techniques Component: FR-S-VIS-002	227
Table 215: Visualisation and Reporting Techniques Component: FR-S-VIS-003	227
Table 216: Visualisation and Reporting Techniques Component: FR-S-VIS-004	227
Table 217: Visualisation and Reporting Techniques Component: FR-S-VIS-005	228
Table 218: Visualisation and Reporting Techniques Component: FR-S-VIS-006	228
Table 219: Visualisation and Reporting Techniques Component: FR-S-VIS-007	228
Table 220: Visualisation and Reporting Techniques Component: FR-S-VIS-008	228
Table 221: Visualisation and Reporting Techniques Component: FR-S-VIS-009	229
Table 222: User Interfaces: NFR-PER-UI-001	229
Table 223: User Interfaces: NFR-PORT-UI-001	229
Table 224: User Interfaces: NFR-SEC-UI-001	229
Table 225: User Interfaces: NFR-UI-UI-001	230
Table 226: Visualisation and Reporting Techniques Component: NFR-COMP-VIS-001	230
Table 227: Visualisation and Reporting Techniques Component: NFR-MAIN-VIS-001	230
Table 228: Visualisation and Reporting Techniques Component: NFR-PER-VIS-001	231
Table 229: Visualisation and Reporting Techniques Component: NFR-REL-VIS-001	231
Table 230: Visualisation and Reporting Techniques Component: NFR-SEC-VIS-001	231
Table 231: Visualisation and Reporting Techniques Component: NFR-UI-VIS-001	231
Table 232: TANGO Integrated platform: NFR-COMP-TIP-001	232
Table 233: TANGO Integrated platform: NFR-MAIN-TIP-001	232
Table 234: TANGO Integrated platform: NFR-MAIN-TIP-001	232
Table 235: TANGO Integrated platform: NFR-PER-TIP-001	233
Table 236: TANGO Integrated platform: NFR-PER-TIP-002	233
Table 237: TANGO Integrated platform: NFR-PER-TIP-003	233
Table 238: TANGO Integrated platform: NFR-PER-TIP-004	234
Table 239: TANGO Integrated platform: NFR-PORT-TIP-001	234
Table 240: TANGO Integrated platform: NFR-PORT-TIP-002	234
Table 241: TANGO Integrated platform: NFR-REL-TIP-001	234
Table 242: TANGO Integrated platform: NFR-COMP-TIP-002	235
Table 243: TANGO Integrated platform: NFR-SEC-TIP-001	235
Table 244: TANGO Integrated platform: NFR-SUST-TIP-001	235
Table 245: TANGO Integrated platform: NFR-SUST-TIP-002	236
Table 246: Legal Requirement LEG-001	236
Table 247: Legal Requirement LEG-002	236
Table 248: Legal Requirement LEG-003	237
Table 249: Legal Requirement LEG-004	237
Table 250: Legal Requirement LEG-005	237

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	13 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

List of Figures

Figure 1: TANGO Methodology roadmap	25
Figure 2: Fides Design overview	30
Figure 3: DRILL data flow	31
Figure 4: DnAMS data flow	32
Figure 5: Fides implementation for KYC/KYB use case	33
Figure 6: Trustworthiness Scoring – Staggered Schema	37
Figure 7: Trustworthiness Scoring – Configuration Process	38
Figure 8: Trustworthiness Scoring – Calculation Components	40
Figure 9: User interest and exchange of Volatile Random Numbers by the UBEM	40
Figure 10: User Recommendations	41
Figure 11: Synchronizing VRNs across vendors	41
Figure 12: CP module architecture	47
Figure 13: CP-ABE workflow	48
Figure 14: CP-ABE Keys Retrieval	48
Figure 15: Self-Encryption workflow	53
Figure 16: Illustration of the individual modules of the SSI component	57
Figure 17: IoT Device SSI component and interwork	58
Figure 18: Seamless user onboarding information flow	63
Figure 19: IoT Device SSI component and its interwork for onboarding	64
Figure 20: Information flow for the different components involved regarding the UCBA component	67
Figure 21: Device Continuous Behavioural Authentication component's architecture	71
Figure 22: Compiler's internal architecture	75
Figure 23: EDAE's High Level Architecture	81
Figure 24: T5.2 Main Components and their Interactions	86
Figure 25: Federated Learning Component High-Level Description	87
Figure 26: MLOps High-Level Depiction	88
Figure 27: AutoML Component Simplified Workflow	89
Figure 28: The TornadoVM Architecture	95
Figure 29: Overview of the Data Flow to Execute with TornadoVM	96
Figure 30: PEC Architecture	98
Figure 31: PAT Architecture	98
Figure 32: XAI Module Architecture	103
Figure 33: RENOPS Architecture and Data Flows	107
Figure 34: MVC Architecture Image Source: [33]	111
Figure 35: Visualizations Architecture	114
Figure 36: Overall Architecture	118
Figure 37: TANGO Connector	120
Figure 38: IDS Usage Control Terms (Image source: Position Paper for Usage Control in IDS [35])	124
Figure 39: Example of an IDS Contract Agreement (Image source: Position Paper for Usage Control in IDS [35])	125
Figure 40: YAML configuration file of a Kubernetes Deployment.	126
Figure 41: ConfigMap option	127
Figure 42: Passing environment variables to containers	127
Figure 43: Sample configuration file using the "ImagePullSecrets" option	128
Figure 44: TANGO Management Connector	129
Figure 45: TANGO App Store	133
Figure 46: Example Workflow (1/3)	136
Figure 47: Example Workflow (2/3)	137
Figure 48: Example Workflow (3/3)	138
Figure 49: Automated software release pipelines and service chain deployment options	145

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	14 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

List of Acronyms

Abbreviation / acronym	Description
ABE	Attribute-Based Encryption
AES	Advanced Encryption Standard
AI	Artificial Intelligence
API	Application Programming Interface
AutoML	Automated Machine Learning
CAV	Connected and Automated Vehicle
CI/CD	Continuous Integration/ Continuous Delivery and/or Deployment
CP-ABE	Cypher text Policy – Attribute Based Encryption
CSV	Comma-Separated Values
CUDA	Programming model and a platform for parallel computing
DCBA	Device-Continuous Behavioural Authentication
DID	Decentralized Identifier
DL	Deep Learning
DnAMS	Decentralized norbloc Access Management System
DNS	Domain Name Server
DoS	Denial of Service
DRILL	Data Redundancy In Legal Limitations - norbloc's decentralized legally compliant data backend
eID	Electronic Identification
eIDAS	Electronic Identification, Authentication and trust Services
EBP	European Blockchain Partnership
EBSI	European Blockchain Services Infrastructure
EC	European Commission
eMRTD	electronic Machine-Readable Travel Document
ESSIF	European Self Sovereign Identity framework
FL	Federated Learning
FPGA	Field Programmable Gate Arrays (integrated circuits)
FR	Functional Requirement
GAIA-X	Initiative created to allow for the secure, open, and transparent use of data
GDPR	General Data Protection Regulation
GPGPU	General-Purpose Graphics Processing Unit
GPU	Graphics Processing Unit
HTTP	Hypertext Transfer Protocol
HIPAA	Health Insurance Portability and Accountability Act
ICAO	International Civil Aviation Organization
IDSA	International Data Spaces Association
IIC	Industrial Internet Consortium
IMEI	International Mobile Equipment Identity

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	15 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

IoT	Internet of Things
IR	Intermediate Representation
JSON	JavaScript Object Notation
JWT	JSON Web Tokens
KA	Key Authority
KPI	Key Performance Indicator
KYB	Know Your Business
KYC	Know Your Customer
MAC	Media Access Control
ML	Machine Learning
MLOps	Machine Learning Operations
NIST	National Institute of Standards and Technology (US)
NFC	Near Field Communication
NFR	Non-Functional Requirement
OIDC	OpenID Connect
OKVS	norbloc's Oblivious Key-Value Storage
P2P	Peer-to-Peer
PET	Privacy-Enhancing Technologies
RF	Radio Frequency
SIOP	Self-Issued OpenID Provider
SSI	Self-Sovereign Identity
SSID	Service Set Identifier
UAM	User Access Management
UBEM	User Behavior Exchange Module
UCBA	User Continuous Behavioural Authentication
UI	User Interface
UPCV	Ubiquitous Personal Context Vector
UX	User Experience
VC	Verifiable Credential
VDR	Verifiable Data Registry
VP	Verifiable Presentation
VRN	Volatile Random Numbers
WLAN	Wireless Local Area Network
W3C	World Wide Web Consortium
XAI	Explainable Artificial Intelligence
ZK	Zero-Knowledge
ZKP	Zero-Knowledge Proof
ZTP	Zero Touch Provisioning

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	16 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Executive Summary

This deliverable encapsulates the outcomes of three major tasks within the project, aimed at laying the foundation for the TANGO framework for transparent, secure, fair and accountable, as well as environmentally sustainable data management. First, a comprehensive overview of the technology offerings encompassed by TANGO (WP3-WP6) is presented, along with details about their envisioned functionality, internal architecture, and interactions with other components. Furthermore, the document illustrates the alignment of these offerings with specific TANGO pilots, as these have been previously specified in D2.2, and outlines the most prominent security threats identified per component alongside the built-in mitigation measures. Additionally, both functional and non-functional requirements (FRs and NFRs) at the system level are provided, reporting also their priority level, associated technology category, and linked user requirements from D2.2. The outcome of this iterative process will serve as a roadmap for enhancing the capabilities of the different components in order to meet the technical specifications set for the TANGO framework.

The initial consolidated version of the TANGO Reference Architecture designed following the IDSA and GAIA-X principles is presented. Divided in three conceptual layers (i.e., trust, management, user) which are distributed across multiple runtime environments and physical locations, the TANGO Reference Architecture leverages multiple containerized services to implement features while enabling the orchestration of the entire ecosystem through a centralized management. The TANGO Connector, the cornerstone of the TANGO architecture serving as the primary entry point to the entire ecosystem is introduced along with the “blueprints” which comprise a set of configuration files for deploying, constructing, and managing the Connectors.

Finally, a preliminary attempt is made towards assessing the Privacy, Ethical, Social, and Legal Impact of TANGO. The results from the performed data mapping process reveal the critical procedures that need to be continuously monitored to ensure GDPR compliance at all times. The set of technical and operational measures that have been implemented to guarantee data security and privacy, protecting this way data subjects’ fundamental rights, is detailed in a dedicated subsection. As the assessment of TANGO’s different impacts will be performed in iterations, the project will be evaluated at different stages of the technical implementation, making sure that the project remains aligned with the applicable legal and ethical standards, and current with the needs of the modern digital society.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	17 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

1 Introduction

1.1 Purpose of the document

The objective of this document is to provide the initial outputs of the following tasks:

- **T2.4 Definition of System Requirements & Technical Specifications:** An overview of all the technology offerings involved in TANGO (WP3-WP6) is provided, covering the following aspects:
 - Description of their envisioned functionality, internal architecture, and interoperability with other components.
 - Description of their association with specific pilots of TANGO, based on the corresponding pilot specifications provided by *T2.3: Use Case Scenarios & KPIs Definition*, through [1].
 - Description of the identified security threats per component, accompanied by mitigation measures that each technology offering will provide.
 - A list of functional and non-functional requirements (FRs and NFRs, respectively) that each technology offering is expected to cover at system level. These requirements are strongly related to the User Requirements per defined TANGO pilot which are provided by *T2.2: User Needs and Requirements for Data Management, Processing & Storage*, through [1].
- **T2.5 Security-and-Privacy-by-design Architecture & API Specifications:** The first consolidated version of the TANGO architecture is provided. This architecture is aligned with the principles and reference architecture of IDSA and GAIA-X associations, adopting some of their main concepts and building blocks. In this framework, the envisioned role and placement of the technology offerings (WP3-WP6) in the TANGO reference architecture according to their technical specifications (output of T2.4) has been largely identified and described.
- **T2.6 Privacy, Ethical, Social and Legal Impact Assessment:** The results of the preliminary comprehensive impact assessment, known as Privacy Impact Assessment Plus (PIA+), covering privacy, ethical, social, and legal issues related to the TANGO framework are provided. As a first step, a data mapping process has been implemented, where partners have been asked to report on their activities, personal data processing actions, data storage, security measures, and possible ethical considerations that may be raised related to data subjects' rights.

1.2 Relation to other project work

As described in section 1.1, the present document is strongly related to the outputs of T2.2 and T2.3, in the context of describing how the TANGO technical offerings will associate with specific pilot use cases and address specific user requirements. Moreover, considering the analysis provided in *T2.1: GAP Analysis in Distributed Data Management, Processing & Storage* [2] regarding the investigation of state-of-the-art within the technology areas of TANGO, the corresponding technical offerings go beyond that in providing detailed descriptions of their envisioned functionalities and capabilities.

Moreover, the present document will serve as the basis for *T6.1: Continuous Integration and Delivery* (starting in M13) that will provide the deployment view and instantiation of the TANGO architecture. Any updates/refinements in the architecture (T2.5) as well as the other WP2 reporting tasks will be provided in *D2.4: System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment Final version* (M28).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	18 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

1.3 Structure of the document

This document is structured in ten major sections, as presented below:

Section 2 provides an overview of the TANGO concept and how it is associated with the TANGO key technological areas, as well as the methodological framework adopted in the project.

Sections 3-6 present in detail the technical offerings which are adopted in TANGO in the technology domains listed below, covering the aspects outlined in section 1.1 in the context of T2.4:

- Distributed Privacy-preserving Data Management and Storage (section 3).
- Distributed Trust Management Framework (section 4).
- AI-based Framework for Green and Trustworthy Operations (section 5).
- Data visualization techniques and User Interfaces (section 6).

Section 7 delivers the first version of the TANGO reference architecture.

Section 8 presents an overview of the TANGO integrated platform, which will incorporate the different technical components into an end-to-end system, as an instantiation of the TANGO reference architecture.

Section 9 documents the initial outcomes of the TANGO impact assessment, covering privacy, ethical, social and legal issues.

Section 10 concludes the document and provides an overview of the next steps.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	19 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

2 Concept & Methodology

2.1 TANGO Concept & Key Technological Areas

TANGO aims to provide a novel solution that ties together several technological offerings. These offerings will be designed, developed, and integrated together in order to enable fair, accountable, trustworthy, and eco-friendly data sharing and management among end users and organizations. The developed TANGO platform will be able to handle large volumes of data originating from heterogeneous sources and perform all the required processing and privacy-enhancing operations, which will allow the sharing with the relevant stakeholders through a secure, sovereign, trusted, and privacy-preserving manner. TANGO will promote decentralization, avoiding dependency on central authorities, by adopting a distributed identity management framework, while it will also foster interoperability through well-defined interactions between the components that constitute its platform, and ensure compliance with the recent and widely adopted paradigms of GAIA-X and IDSA.

The TANGO platform will be built around three main pillars that also represent the three key technical objectives of the project. The description of the key technological areas that will be employed and the technology offerings which will be implemented to achieve these objectives is presented in the sections below:

2.1.1 Distributed Privacy-preserving Data Management and Storage

To achieve this objective TANGO will leverage a combination of technological areas and offerings that focus on ensuring privacy, security and accountability of data in a decentralized and distributed manner.

Technological Areas & Offerings:

- **Blockchain-based Data Storage and Sharing:** A distributed data model that is based on Blockchain, increases the degree of security and performance efficiency, and provides significant advantages for improved data management and sharing: increased level of security, untampered data streams and integrity and increased level of trust amongst stakeholders.

TANGO Tech Offering: The project partner NOR will provide its data sharing platform Fides, which is a distributed, privacy-preserving, fair, accountable data storage and sharing platform that enables the exchange of highly confidential data in a secure and sovereign manner. Fides allows the sharing of regulated data coupled with its attestation/validation stamps and in an environment that records immutably all actions, most importantly customer consent.

- **Trustworthy Data Sharing:** The term ‘trustworthiness’ is used to describe the quality of existing and future relationships between companies, people, systems, and components. A trustworthy system ensures that all its components behave in an expected manner. For supply/value chain security and risk management, the term ‘trustworthiness’ corresponds to the supplier’s ability to meet the expectations of the potential contract partner in a verifiable way.

TANGO Tech Offering: Trustworthy Data Sharing module will provide a set of tools supporting access and usage control of data. Three major components will be developed: a) a Data Usage Control component providing a Privacy Risk Score, b) a Trustworthiness model for calculating a Trustworthiness Score, c) a privacy preserving recommendation module using Ubiquitous Personal Context Vectors (UPCVs).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	20 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ **Confidentiality and Privacy by Design:** Data-centric approaches tackle security and privacy in a holistic manner. Privacy by design principles include proactive protection instead of remedial action when violations happen, privacy as a default setting, privacy embedded into the design, full functionality with privacy protection, privacy protection through the complete lifecycle of the data, visibility and transparency, and respect for user privacy.

TANGO Tech Offering: The confidentiality and authentication of shared data will be supported by a module for creation and application of sticky policies and signatures (SP). Data will be encrypted according to an attribute-based policy set by the data owner. Potential consumers will need to retrieve keys associated to their identity attributes and will only be able to decrypt data if their attributes (i.e., the respective keys) fulfil the sticky policy. Policy-based access control will be used to address user consent (UC). Attribute-based policies will be used to enable fine-granular checks based on identity attributes.

- ▶ **Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms:** Self-encryption is a method of data protection in which the device or system where the data is stored is responsible for both encrypting and decrypting the data, rather than relying on a separate system or software to do so. Self-encryption was introduced as a method of encrypting files that does not require user intervention or passwords.

TANGO Tech Offering: Self-encryption and decryption will be used as efficient methods to protect data generated by different use cases. Multi-factor information recovery supports secure access to user and other data on cloud servers if audit services are required. Only individuals who provide and match multiple factors can access sensitive information on the cloud server and thus reducing the risk of data theft or misuse.

2.1.2 Distributed Trust Management Framework

This objective underpins the secure access and interaction of users and devices within the TANGO platform.

Technological Areas & Offerings

- ▶ **Self-Sovereign Identity Management:** Self-sovereign identity (SSI) approach improves user authentication with added privacy and trust and empowers end-users with more control over their own personal data. The concept of SSI is based on the interaction between the actors Issuer (issuing credentials), Holder (holding the credentials, e.g., in their wallet application), and Verifier (verifying the credentials), in addition to the Data registry (e.g. Blockchain) that includes the functions to verify the credentials of the user. SSI empowers individuals by improving their right to self-determination regarding their personal data, providing interoperable infrastructure for human-centric personal data management and governance.

TANGO Tech Offering: TANGO will offer an SSI solution based on Hyperledger Aries GO Framework and Atos Ledger uSelf, which offers high level of assurance supporting dematerialisation of identity and aims at high usability, interoperability and trust based on decentralized infrastructure (Blockchain)

- ▶ **Seamless Onboarding for Users and Devices:** End-users do not trust organizations such as a marketing agency to handle sensitive data e.g. a copy of their passport. Thus, a coherent solution should have the appropriate infrastructure to handle the identity verification of the end-user, without the need for the organization to worry about how to store, process, secure and maintain sensitive personal data such as proof-of-identity. It is essential to provide the end-user the appropriate level of trust regarding the management of personal data. On the other side, in order to create a secure on boarding to a particular service, there is a need to on-board both the user and the device through which a person is accessing the particular service.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	21 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

TANGO Tech Offering: The Onboard SSI will be a mobile remote digital onboarding platform for users that will be able to provide their physical identity such as passport or eID, in order to verify their identity and then forward that information to the SSI management component in order to create the relevant Verifiable Credentials. These credentials will then be accessed in to order to authenticate to different services. The solution will include artificial intelligence components allowing the system to read through an image the information of the identity document, as well as to crosscheck the facial biometric characteristics of the person performing the onboarding, of the person visible on the identity document and of the person on the NFC chip of the identity document. IoT Device Onboarding will provide Zero Touch Provisioning (ZTP) of the devices in that when devices are powered on the SSI component can request directly a public Decentralized Identifier published on the blockchain.

- ▶ **User continuous behavioural authentication:** Authentication constitutes an important component of online and offline systems, ensuring the access control. The primarily purpose is to ensure that the person trying to access a particular system is exactly who they say they are, safeguarding the security of the private data of the system for the particular user, but also preventing any potential cases of fraud.

TANGO Tech Offering: The TANGO solution will use biometrics and behavioural patterns from people's daily life to identify individuals, based on location, interaction, speed, movement, walking, smartphone usage and so on. The solution will continuously authenticate the user, without the need for any user input, by learning the users' behavioural patterns. The authentication process is based on understanding and learning the behavioural patterns of the users by leveraging data extracted from smartphones, such as sensor and device usage data. The system will combine multiple behavioural observations from which the identity of a user can be determined.

- ▶ **Device Continuous Behavioural Authentication:** Continuous Authentication refers to methods that have been used to provide increased security and reliability considering authentication in comparison to traditional authentication methods relying on passwords and certifications. For small IoT devices that feature sufficient computing power to run applications but lack conventional interfaces such as keyboards, touchscreens or sensors such as gyroscope and accelerometer, conventional continuous authentication and authorization methods that exploit user behavioural patterns collected by embedded sensors cannot be applied. As such, device continuous behavioural authentication is a method of authenticating devices based on the way those devices operate and interact with the network and ambient environment.

TANGO Tech Offering: The offered component will provide all the required functionality for the continuous authentication of a user's device via the online behavioural analysis and learning of the unique patterns and characteristics of a device's operation such as its hardware configuration, software settings, usage patterns, and network activity. Through the analysis of the aforementioned patterns and characteristics, a baseline is determined, and a continuous monitoring mechanism compares it to the established baseline.

- ▶ **Hardening Against Side-channel Attacks:** Side-channel attacks represent a dangerous threat against embedded systems and IoT devices as they are efficient at finding encryption keys, in particular for the AES block-cipher. Countermeasures against side-channel attack are divided in 2 categories: a) masking countermeasures, which make attacks more complicated and b) hiding countermeasures that consist in lowering the signal-to-noise ratio.

TANGO Tech Offering: The provided component will be a compiler able to automatically apply side-channel attack countermeasures. It will consider code polymorphism and loop shuffling countermeasures.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	22 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

2.1.3 AI-based Framework for Green & Trustworthy Operations

TANGO will emphasize environmental sustainability and responsible operations, driven by artificial intelligence.

Technological Areas & Offerings

- **Exploratory Data Analysis Engine:** Exploratory Data Analysis builds upon the soundness of statistical models and hypothesis formulation, which are also used in the classical approach, to reveal hidden and unknown information from data in a form that enables analysts to obtain an immediate, direct and easy-to-understand representation of it.

TANGO Tech Offering: The exploratory data analysis engine (EDAE) will be fed with data deriving from the end users aiming to perform data analytics with a focus on the identification of patterns or features of interest among the data. EDAE will be capable of: a) identifying correlations among the heterogeneous real-world data that will be inserted into the system and b) selecting correlation-based features. A combination of well-established data analysis techniques will be exploited in order to develop an AI-driven method to automatically perform the correlation analysis, visualising the generated results.

- **Energy Efficient AI Model Training:** AutoML stands for Automatic Machine Learning and includes methods to build and validate ML pipelines minimizing the intervention of a user, which is of great interest for final users, such as enterprises. AutoML is expensive in terms of resource consumption, as many approaches rely on large experiments over many datasets, especially while benchmarking many of the existing solutions.

Similar to the concept of DevOps for shortening and automating the system development, delivery and deployment, the concept of MLOps tries to do the same for the entire life cycle of the Machine Learning models management.

Federated Learning (FL) aims to train Machine Learning models with data distributed among across multiple devices and computing nodes without the need of disclosure them. It is therefore a technique that brings the ML algorithms to the data and then combines the partial training results in a new improved trained model.

TANGO Tech Offering: There are three components which will be developed to cover the concepts of the Energy Efficient AI Model Training area in TANGO.

- The **MLOps** component will enable the management of the different ML models to be created in TANGO. This management will cover the whole lifecycle of the models, from their conception to the training, serving, and monitoring phase. It is intended to be used by ML and AI developers and AI system engineers.
 - The **AutoML** component will provide support for the automation of repetitive steps to design and train machine learning models. This component will help regarding technical issues involved in the process of setting up and designing predictive models (such as hyperparameter tuning or architecture selection in case of neural networks), aiming for an automatic optimization of the model to be created.
 - The **Federated Learning (FL)** component aims at providing the functionality of Federated Learning (FL) in TANGO. By using Federated Learning techniques, models can be generated from different datasets without the need of the sharing of these datasets. This can be quite useful in scenarios where predictions based on datasets that involve sensitive data or are linked to policies that do not allow their sharing are needed.
- **Dynamic Intelligent Execution on Heterogeneous Systems:** Regarding Java and managed programming languages, prior work has focused on embedded support for heterogeneous programming inside existing languages targeting GPGPUs, FPGAs, vector units, and multi-core processors. The majority of these efforts rely on translating bytecode into CUDA or OpenCL C to generate code for the GPGPU.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	23 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

TANGO Tech Offering: TANGO will employ TornadoVM, an open-source programming framework that enables hardware acceleration for managed programming languages. TornadoVM enables hardware acceleration on a single node for data processing parallelization and plans execution on hardware accelerators based on energy efficiency. TornadoVM will be integrated with Exploratory Data Analysis Engine to accelerate efficiently the data processing.

- ▶ **Privacy Threat Modelling and Identification for Trustworthy AI:** A privacy risk assessment, also known as data protection impact assessment or privacy impact assessment, is the process of identifying and evaluating privacy risks. Better management of privacy risks through the development of effective solutions to protect individuals' privacy when designing or deploying systems, products and services that process data can help organizations build customer trust.

TANGO Tech Offering: TANGO will offer two components for privacy impact assessment and privacy assurance:

- The **Privacy Enhancing Component (PEC)** which aims to provide a tool, based on the NIST Privacy impact assessment methodology, to perform privacy impact assessment, and
 - The **Privacy Assurance Tool (PAT)** which aims to monitor privacy related to the consent and data policy between data subjects and organisations.
- ▶ **X-AI for Privacy and Trust Enhancement:** Explainable AI (X-AI) is a collection of complex approaches designed to make the outcomes of a solution intelligible to humans, for the purpose of validating real AI use cases. It stands in stark contrast to the notion of the "black box" in machine learning (ML), where even the developers or designers may be unable to explain why an AI model gets a specific conclusion. In terms of security, the false predictions might make the system vulnerable to attack and lead to the implementation of zero-trust security for vital systems.

TANGO Tech Offering: The XAI module will be used to describe an AI model, its expected impact and potential biases. It will help characterize model accuracy, fairness, transparency, and outcomes in AI-powered decision making. Apart from helping build trust in the predictions the ML models produce, it will also help debug the models in cases where bias in the decision making process is discovered.

- ▶ **Infrastructure Management Based on AI:** A novel approach for infrastructure management by using the federated AI approach and balancing the trade-off among system responsiveness, energy-consumption, privacy and data-protection is needed.

TANGO Tech Offering: Through TANGO, we would introduce RENOPS (Renewable Energy Forecast Production Service). A service that provides information on renewable energy potential via an easily accessible API. This information will be used to make data-driven decisions such as scheduling of energy-intensive tasks. In overall, the solution would decrease the carbon footprint of energy-intensive processing tasks promoting a green and energy efficient approach.

2.1.4 TANGO Integrated Platform Concepts

The TANGO integrated platform will consist of all the aforementioned technology offerings and services that will be integrated into an end-to-end system. TANGO will use a CI/CD system that follows all DevOps best practices for software development, versioning, testing and deployment and builds automated pipelines for testing and deployment. Additionally, several supportive technologies will be developed that enable the interaction of human users with the TANGO platform and which consist of user interfaces and visualization techniques. A web-based interface will be provided that incorporates all the options for the end-user and/or the administrator of the TANGO platform, to access and parametrize the various TANGO

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	24 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

offerings, monitor their usage, manage users, and see statistics and information about their usage. Furthermore, visualisation and reporting techniques will be designed and developed to analyse large amounts of data and allow administrators for more informed and data-driven decisions.

Overall, TANGO will provide a comprehensive, eco-friendly, and trustworthy solution to overcome the modern challenges of data management and sharing.

2.2 Methodological Framework

2.2.1 Project Methodology

As described in the DoA, the project methodology has been based upon the identification of the operational challenges that stakeholders face in the context of data operations in data spaces, across the data life cycle. The TANGO methodology follows a 4-step approach, where each of these is associated with a project Work Package (WP) that includes the tasks planned to address the project objectives as presented in Figure 1. The *Baseline* step has been triggered in WP2 by the in-depth investigation of the current state-of-the-art for all prominent technologies leveraged in TANGO (D2.1), followed by the elicitation of the user needs & requirements and the definition of the use case scenarios (D2.2). Deliverable D2.3 concludes this set of essential deliverables produced as part of the *Baseline* step, reporting on the system requirements and technical specifications of all technical offerings, delivering TANGO's reference architecture designed in line with the IDSA/GAIA-X standards to ensure compatibility, and finally, identifying relevant ethical, legal, social and privacy related themes, data flows, and possible associated risks.

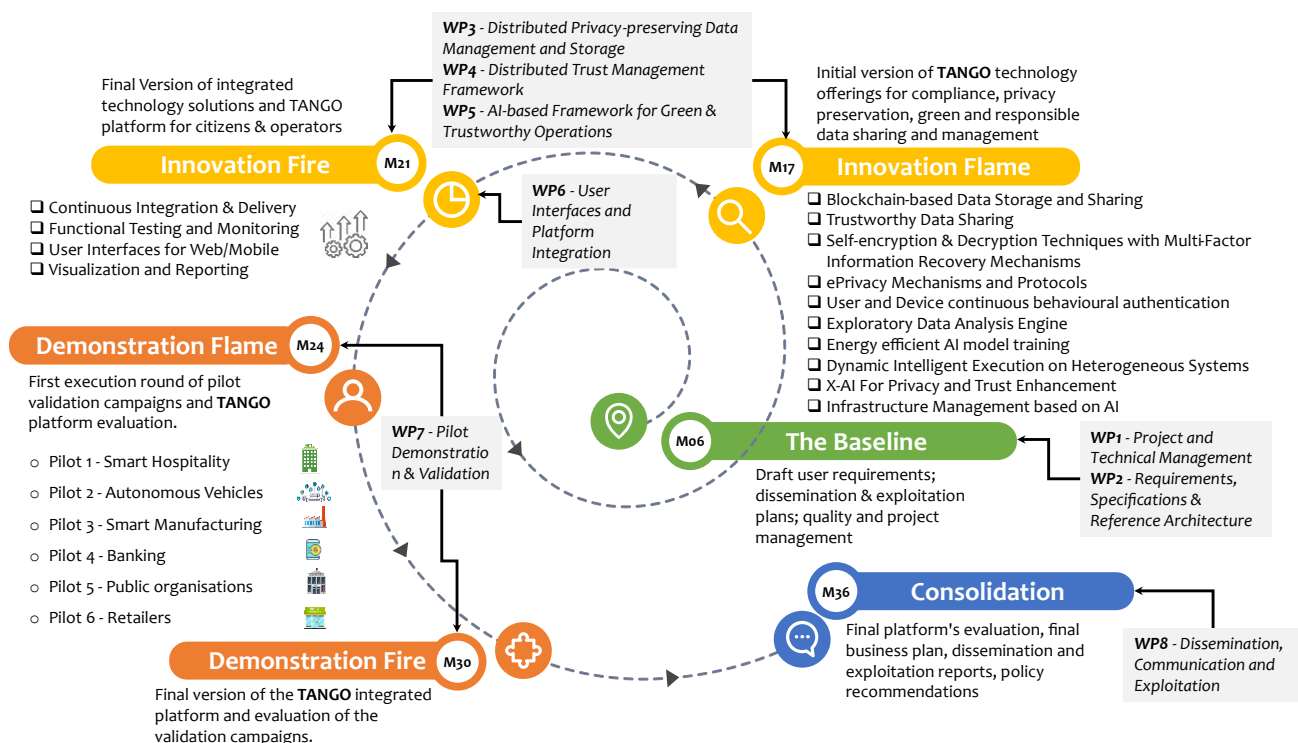


Figure 1: TANGO Methodology roadmap

In the next phase, known as the "Innovation Flame," we will leverage the insights extracted from the previously mentioned deliverables as a foundation for the technical Work Packages

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	25 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

(WPs) to integrate their individual technologies within an agile, decentralized, and highly scalable architectural framework. This framework is designed to facilitate the seamless coordination of various services operating within the broader TANGO ecosystem. The elicited system requirements will serve as a roadmap for enhancing the capabilities of the different components, aligning them with the precise technical specifications set for the platform. This process essentially involves expanding the functionalities of each component to ensure they align seamlessly with the platform's defined technical requirements and objectives. Finally, the privacy, ethical, and legal principles defined in this deliverable will serve as a compass for the technical partners to seek and incorporate appropriate technical or operational measures with special attention given to privacy, social impact, security, lawful basis, and aspects of surveillance.

2.2.2 Requirements Methodology

The system requirements definition is an essential step of the Baseline phase for the success of the TANGO project. Driven by the outcomes of T2.2 and T2.3 which are documented in D2.2, the definition of system requirements aims to describe how TANGO will be implemented to provide the necessary functionalities and attributes for fulfilling the expressed user requirements. In addition, an initial identification of technologies and threats is performed for each of the offerings. This information included in this deliverable can then be utilized in the next phase to guide the development efforts of the TANGO platform.

2.2.2.1 Process Definition

To achieve the goal of the collection of requirements and technical specifications, the INTRA team had to design a strategic approach for collecting and managing inputs from the technical partners. In the case of the TANGO project, this process becomes more challenging due to the complexity of the system and the involvement of numerous partners with different technical offerings. The process involved the design of the deliverable's structure, the collection of inputs in three phases, and the collaboration with the partners responsible for the user requirements and use case descriptions.

The first step of the requirements elicitation was the definition of the table of contents of the deliverable, to serve as a blueprint for the information that should be gathered. Then, to efficiently gather inputs from the technical partners, an iterative approach was followed. The foundational step to gathering initial inputs was the creation and distribution of a "Requirements Collection Questionnaire", one for each technical offering. Once the completed questionnaires were received, the INTRA team consolidated the responses into a single document, carefully reviewed the contents to identify potential gaps, ambiguities, or inconsistencies, and informed the partners to refine their inputs in the next round of input collection. During this process, INTRA cooperated closely with the UPRC and LSTECH teams which were handling the use case and user requirements definition (D2.2) to ensure that the system requirements were clear, accurate, and aligned with the TANGO objectives. These efforts result in the content presented in the deliverable sections 3 "Distributed Privacy-preserving Data Management and Storage" to 6 "Visualisation Techniques & User Interfaces". Of course, it is expected that this information will not remain static but will likely evolve and change during the lifetime of the project and as the software development tasks progress, following the modern practices of Agile development. Their level of completion will also be assessed and validated, via functional and integration testing (defined in the context of WP6) and through the pilot demonstrations (WP7).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	26 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

2.2.2.2 Format & terminology / Requirements Classification

The term “System Requirements” includes two distinct types of requirements: Functional and Non-Functional, each of which play an important role in determining the characteristics and behaviour of the system.

Functional requirements (FRs), as the name implies, are used to describe the specific functions and features that the system should incorporate to meet user needs. It is important to prioritize these functional requirements to maximize efficiency and ensure that developers are focused on delivering the most critical features first. For this reason, the “MoSCoW” method [3] was employed when describing functional requirements, and an extra field was added to explicitly define the priority (where the keywords ‘must’, ‘should’, ‘could’ mapping to high, medium, and low priority respectively). This approach was followed to distinguish which features are most important, which can be considered, and which would be nice to have but not necessary in a clear, structured manner. In addition, to organize the large number of requirements which was expected for a system as complex as TANGO, a Technology Category Field was assigned to each functional requirement, based on the corresponding WP(s) the requirements are addressed in. The categories created are: “DATA”, “TRUST”, “AI”, and “SUPPORT”, corresponding to WPs 3,4,5, and 6, respectively. Also, since functional requirements are the translation of user requirements from a technical point of view, it is necessary to highlight this mapping between them with a dedicated field by using the ID codes of assigned in deliverable D2.2. Finally, a field explaining the rationale of the functional requirement was added. All information identified as essential for fully describing a functional requirement is included in the final table of information that was requested to be completed for each functional requirement the technical partners expressed.

Table 1: Template for each Functional Requirement description

ID	<FRCODE>* To be filled by Author	<PARTNER>
Name	Please provide a short and comprehensive name for the requirement (in a sentence or less).	
Description	The TANGO system must/should/could ...	
Priority	HIGH / MEDIUM / LOW	
Technology Category	DATA (corresponds to WP3) / TRUST (corresponds to WP4) / AI (corresponds to WP5) / SUPPORT (corresponds to WP6)	
Related User Requirement	If applicable, please include the ID of the corresponding User Requirement (available in D2.2).	
Rationale	Please state the rationale behind including this requirement. e.g., This feature will guarantee that only authorised users will access the platform, which is necessary according to the data security and privacy principles.	

Non-functional requirements (NFRs) involve qualitative properties that do not refer to a specific feature of the system, but rather to attributes of its whole operation. To assist the partners during their collection, and to provide a way of organizing them, the INTRA team defined a set of specific categories for the non-functional requirements, which are presented in the table below. The below categories have been defined taking into account the project needs, the ISO/IEC 25010:201 [4] list of non-functional properties for systems and software engineering, and the well-known taxonomy of 27 useful non-functional requirements that

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	27 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

address adaptation, design, sustainment and viability concerns during systems design endeavours [5].

Considering the growing importance of environmental sustainability for organizations, driven by concerns from customers, employees, and stakeholders, more and more stakeholders are now developing strategies for sustainable operations, including net-zero emissions and lower greenhouse gas targets. Sustainability has created the need for new and adjusted business capabilities with additional requirements around data and processes that consider energy efficiency. Recently, there's been an increasing number of sources proposing the definition of sustainability as a non-functional requirement due to the inherent need to integrate it into the entire development lifecycle, as well as into the deployment process [6], [7], [8]. By including sustainability as a horizontal NFR, the solution providers can ensure that their technologies align with sustainability principles from the beginning.

To this end, for the purposes of TANGO we have extended the list of traditional NFR categories by adding the NFR_SUST category concerning the qualities that describe the capability of a software solution to be energy efficient, mitigating its ecological footprint during deployment and execution.

Table 2: Defined Categories of Non-Functional Requirements

Code	Category	Description
NFR_COMP	Compatibility	Qualities that describe the degree to which the software can exchange information with other systems. Includes the concepts of co-existence and interoperability
NFR_MAIN	Maintainability	Qualities that describe the ability of the software to be changed, enhanced and supported over time with a fair degree of effortlessness. Includes the concepts of modularity , reusability , modifiability etc.
NFR_PER	Performance	Qualities that describe the relationship between the level of performance of the software and the number of resources used, under stated conditions. Includes the concepts of capacity , resource utilization , and scalability .
NFR_PORT	Portability	Qualities that describe the ability of the software to be transferred from one environment to another. Includes the concepts of adaptability , replaceability etc.
NFR_REL	Reliability	Qualities that describe the capability of the software to maintain its level of performance under stated conditions for a stated time period. Includes the concepts of availability , recoverability , fault tolerance etc.
NFR_SEC	Security	Qualities that describe the protection of the software against attacks. Includes the concepts of confidentiality , integrity , authenticity , accountability etc.
NFR_UI	Usability	Qualities that describe the effort needed by the users to use the software. Includes the concepts of accessibility , user error protection , UX , UI aesthetics etc.
NFR_EXIST	HW and SW Constraints	Hardware and software constraints due to integration of existing assets.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	28 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

NFR_SUST	Sustainability	Qualities that describe the capability of a software solution to be energy efficient, mitigating its ecological footprint.
----------	----------------	---

In addition to these categories, partners leading technical components were asked to provide a priority for each of the non-functional requirements, following the same conventions as the priorities defined in the functional requirements, to associate them with user requirements where possible, and to provide a rationale for their selection. All system requirements (functional and non-functional) were assigned to a unique ID for easier identification/tracking. For the reader's convenience and for brevity's sake, the tables including the full details of the system requirements have been placed in the Annex (A. System Requirements for Distributed Privacy-preserving Data Management and Storage-F. Privacy, Ethics, Legal, and Societal Requirements), and summary tables have been included in the corresponding sections within the main body of the document.

2.2.3 Architecture Methodology

In order to bring together all the technology offerings that TANGO consortium offers and integrate them into a coherent, flexible and efficient platform, a reference architecture has been carefully designed by considering the key technical objectives of the project, the specifications and functional requirements of the components, the needs of the use-cases and the contemporary widely adopted paradigms for trustworthy data sharing and management of GAIA-X and IDSA (section 7.1.2). The architecture spans over three layers that provide Trust, Management and User functionalities, with the components from all layers closely interacting with each other. It follows the microservices approach that offers significant benefits such as scalability and flexibility, resilience, better resource utilization and improved security, all of which contribute to the project's methodology for achieving its objectives.

The key notion of the architecture is the Connector (Section 7.2), which is essentially the point of entry for a Person, Organization or IoT device to the TANGO ecosystem. It consists of components (services) from all layers and provides the means for isolation and secure and trustworthy data exchange between stakeholders. Connectors highlight the decentralized concept of TANGO by allowing mutual authentication and authorization without the need for a centralized authority, by incorporating the SSI solution. Data transactions are completely transparent with the employment of logging and clearing mechanisms and usage control modules inside the Connector enforce data usage according to the terms of the data Provider.

The TANGO reference architecture is highly modular, thanks to the microservices approach, and allows for high configurability and flexibility. In more detail, the architecture enables the users and administrators to configure the TANGO system and tailor it to their needs by defining the notion of a “blueprint”, which is essentially a configuration file that can configure each component/technology offering and specify the interactions and inter-dependencies between them. As every use-case has its specific needs and requirements, a corresponding “blueprint” will re-configure/re-instantiate the platform by enabling and configuring a subset of components that are needed for this specific use-case. As such, the architecture is designed to offer a fully-fledged system, but which, at the same time, can effortlessly adapt to the needs of a use-case.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	29 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

3 Distributed Privacy-preserving Data Management and Storage

3.1 Blockchain-based Data Storage and Sharing [T3.1]

3.1.1 Description & Functionality

The Fides platform offers decentralized storage of regulated data / attestations together with verification stamps and sharing stored data across interested parties.

The full solution is truly decentralized and offers legal compliance, great data redundancy, system resilience and enforced data verification upon entry with support for independent data verification entities.

Fides is a holistic solution which can be implemented independently or can be used as a part of another project. The latter is the case in the realm of TANGO project; in this case, some of the internal components of Fides will not be fully utilized as their functions will be either redundant or irrelevant to the goals of TANGO. Every such case of reduced utilization is indicated in the description of the respective internal Fides component.

3.1.1.1 Fides internal components

Figure 2 presents an overview of the Fides design and its internal components.

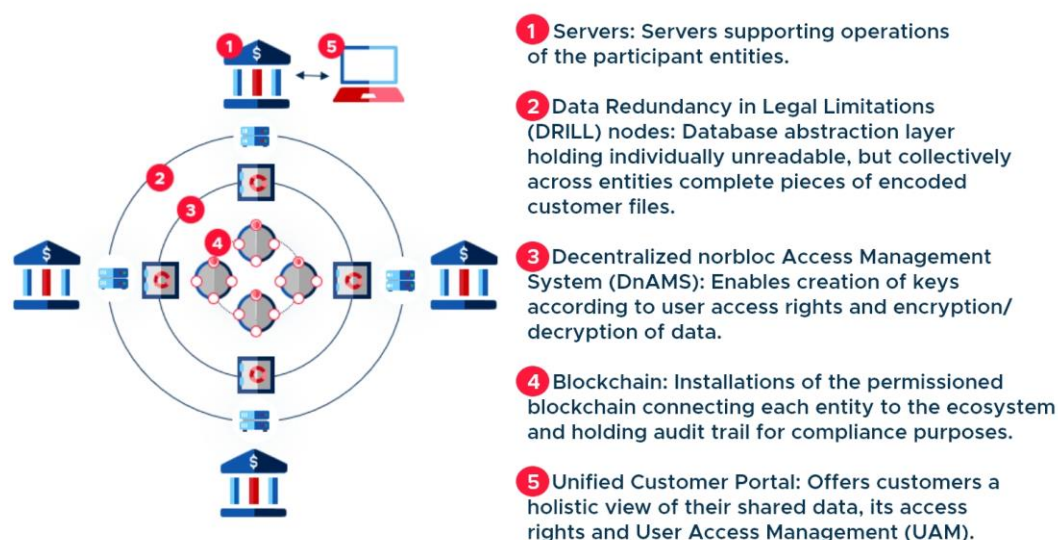


Figure 2: Fides Design overview

The Fides internal components that will be utilized in TANGO are:

- ▶ Proprietary norbloc data backend: DRILL – fully utilized in TANGO
- ▶ Proprietary norbloc access management system: DnAMS – partially utilized in TANGO
- ▶ Blockchain (any private blockchain can be used). By default, Fides offers Hyperledger Fabric blockchain – fully utilized in TANGO

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	30 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

3.1.1.1.1 Internal component: DRILL

In essence DRILL is a peer-to-peer (P2P) network comprising DRILL nodes, which are an essential part of every Fides node. DRILL is designed to store individual data pieces in an environment where conformance to privacy regulations is an essential requirement. To facilitate privacy regulations, DRILL includes symmetric data encryption as a part of its design.

Figure 3 describes the main steps of DRILL data flow. Each piece of data is independently encrypted, split in several chunks and those chunks are distributed between DRILL peers in such a way that no one of the peers stores all chunks. Therefore no one of the peers is storing private data from a GDPR perspective, as individual chunks are unintelligible.

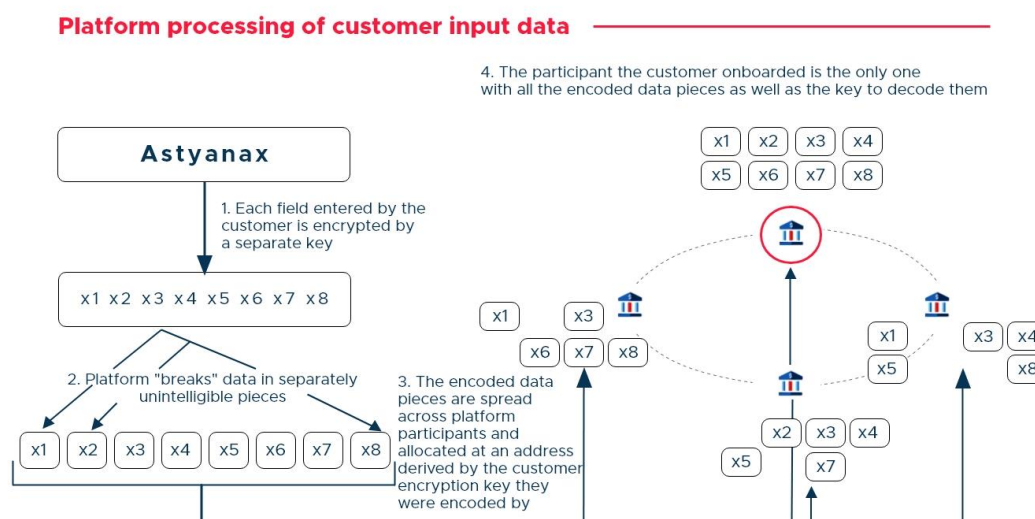


Figure 3: DRILL data flow

When DnAMS component is presented with a “data access permission proof” it releases the necessary information from blockchain storage to a DRILL peer, which allows the DRILL peer to retrieve all chunks from other peers and reconstruct the full encrypted data. Then DnAMS component on the basis of the same “data access permission proof” releases the encryption key to decrypt the reconstructed encrypted data.

3.1.1.1.2 Internal component: DnAMS

DnAMS is a decentralized access management system. It manages symmetric data keys used by DRILL encryption of individual data pieces; also, it manages private keys of data subjects, thus removing a need in maintaining separate external wallets. DnAMS manages access rights by managing respective encryption keys. DnAMS allows release of a key only upon being presented with a valid “data access permission proof” (“proof of consent”).

Figure 4 describes the main steps of DnAMS data flow.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment	Page:	31 of 237
Reference:	D2.3	Dissemination:	PU
	Version:	2.0	Status:
			Final

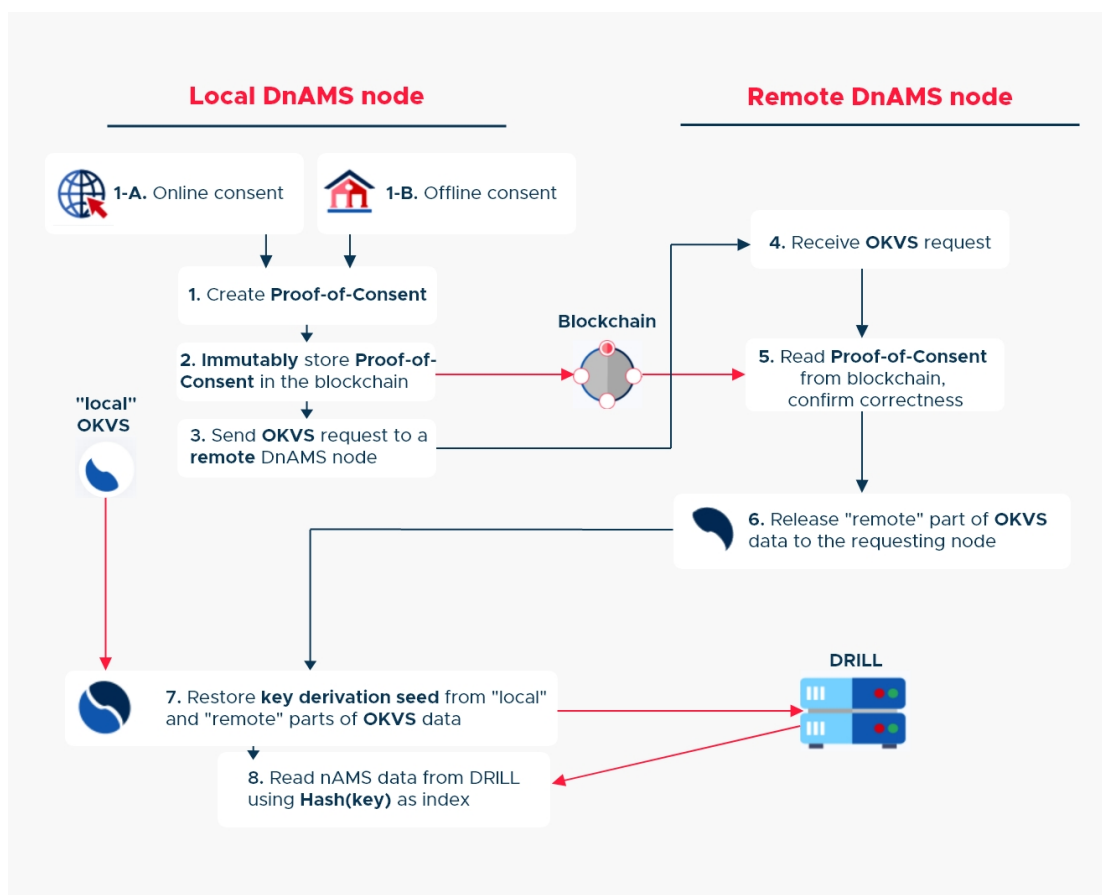


Figure 4: DnAMS data flow

Features that distinguish DnAMS from other key management/access management systems are:

- Explicit consent management (see Figure 4).
- True decentralization – all DnAMS node are of the same rank, all of them together form a flat P2P DnAMS network.
- Local obliviousness: By exploring only local data storage it is impossible to deduce any knowledge about keys or data stored, including even simple boolean fact if a certain user data is present in the system or not. This is achieved by utilizing norbloc proprietary technology OKVS – Oblivious Key-Value Storage.

3.1.1.1.3 Internal component: Blockchain

A blockchain is a helper component that brings in such properties as immutable storage, immutable audit log and reliable timestamping. It is heavily used by other components; DnAMS and DRILL:

- DRILL stores in blockchain data hashes, thus ensuring data integrity;
- DnAMS stores in blockchain “proofs of consent” / “data access permission proofs”, preventing certain Fides participants from manipulating the proofs to their advantage.

3.1.1.2 Fides Functionality

Fides functionalities that are exploitable by TANGO project include:

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	32 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ **Explicit consent management.** This functionality is delivered by DnAMS internal component, see section 3.1.1.1.2, In the scope of TANGO project other components requiring reliable decentralized data storage will be issuing “proofs of consent” / “data access permission proofs” to facilitate data sharing to the data consumers.
- ▶ **Private data sharing in regulated environments.** This functionality is delivered by Fides design. TANGO project will include Fides component for the use cases that require data sharing in regulated environments and include two or more independent parties. In other use cases where there are no strict privacy regulations, or the data pieces being shared do not constitute private data, or there is only one independent party – other TANGO components shall be used.
- ▶ **Immutable audit logs.** This functionality is delivered by blockchain Fides component. A set of metadata for every data access is immutable stored, ready for external/internal audit.

3.1.1.3 Fides Internal Architecture

Figure 5 describes a Fides-based ecosystem with banks sharing KYC data. In the context of TANGO project Fides will be utilized to store any kind of regulated data that requires sharing, not limited to KYC data only. The exact nature of the data is highly dependent on the use case and specifics of other TANGO components using Fides.

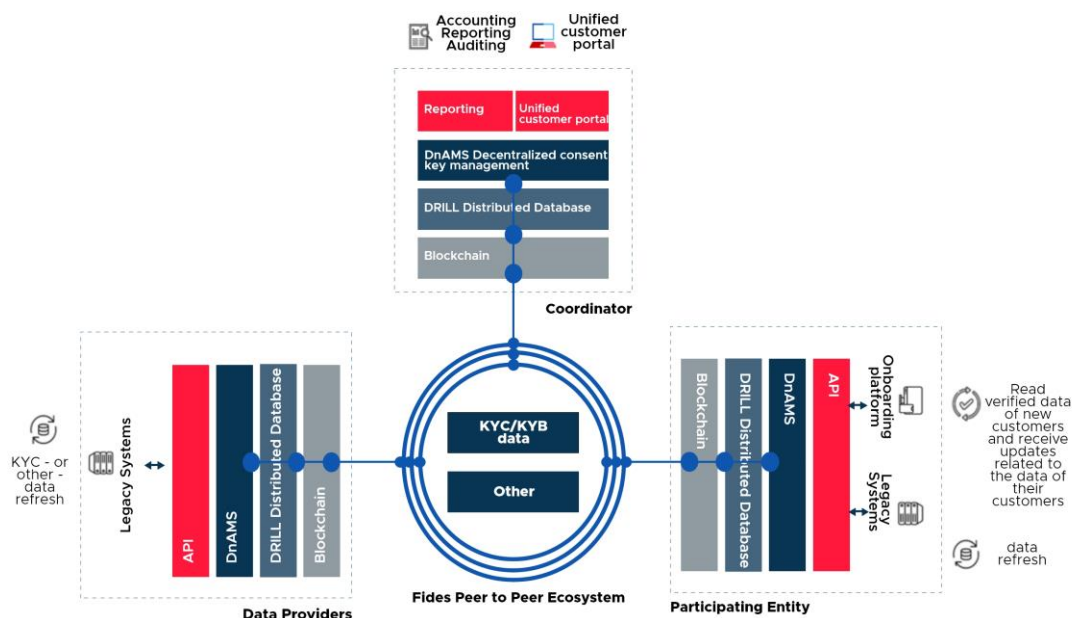


Figure 5: Fides implementation for KYC/KYB use case

To ensure the property of true decentralization and eliminate the danger of having a single central point of control, each participant of Fides platform will have their own installation of a Fides node at their physical or virtual (cloud) premises. At the very least a participant should legally own the installation (for example to own a subscription with the cloud service).

Fides platform will be instantiated as a set of Fides nodes. These nodes will be installed at each participant’s premises. For the case of TANGO pilots, it will be considered to install the nodes at the hosting environment provided by INTRA TANGO partner. The Fides nodes will be communicating with each other, thus creating three peer-to-peer (P2P) networks shown in Figure 5:

- ▶ DRILL P2P: decentralized distributed legally compliant data backend

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	33 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

- ▶ DnAMS P2P: decentralized access management system
- ▶ Blockchain P2P: decentralized blockchain – a glue bringing properties of immutability and reliable timestamping to the system.

From the above description it should be obvious that the crucial prerequisite to implement Fides as a part of TANGO solution is to have at least two independent party in a use case. Fides capabilities cannot be exploited in case TANGO is used by a single party only.

3.1.1.4 Inputs, Outputs, and Interactions with Other Components

Fides component brings in ability to store and share regulated data in presence of strict privacy constraints. Therefore, Fides shall be used by TANGO platform in those use cases that require sharing of the regulated data. If a use case does not have strict privacy regulations, or the data pieces being shared do not constitute regulated data, or there is only one independent party – other methods of data sharing shall be utilized by TANGO platform, as in these use cases Fides will incur significant resource demand without adding any value.

If a usage of Fides is justified by a use case, Fides component will be used as a decentralized data storage and sharing platform by other components requiring such storage and/or sharing. For other components to use Fides, two sets of APIs will be developed:

1. Data API. A set of APIs to be utilized by other components to store data in Fides and request data from in Fides DRILL;
2. Consent API. A set of APIs to be utilized by other components to manage data access permissions in Fides DnAMS.

The exact format of the data stored is not limited by Fides and can be made compatible with IDS/GAIA-X/DID specifications.

The following components may benefit from using Fides as a decentralized data storage and sharing platform:

- Self-Sovereign Identity (T4.1).
- Trustworthy Data Sharing (T3.2).

Furthermore, for the use cases where TANGO users are using decentralized data storage and sharing directly, a proper user interface shall be delivered by User Interfaces component (T6.3).

3.1.2 Associated Pilot Use Cases

Fides shall be used by TANGO platform in those use cases that require sharing of the regulated data. From all the current use case only Pilot 5 (T7.6 – Public organizations) was identified as having such a requirement. Other pilot use cases either do not require data sharing at all, or have only one independent entity, or are sharing non-regulated data.

Table 3: Association of Pilot Use Cases with Fides component T3.1

Pilot	Association
Public organizations	Currently, sharing the data with governmental authorities from Visar and obtaining an application decision from the authorities is a manual and laborious process, prone to errors. Data sharing component aims to improve the situation by automating the data sharing process between applicants, Visar personnel, and governmental authorities, at the same time keeping legal compliance with strict regulations regarding sharing of private data.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	34 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Banking	If there is a need for sharing verified regulated data within the Banking pilot, norbloc's decentralized solution Fides will allow for compliant and secure data sharing.
----------------	---

3.1.3 Threat Landscape

An overview of the threat landscape can be summarized as per the below:

1. **Data Privacy Violations:** Unauthorized access or data breaches can lead to the exposure of sensitive, regulated data, resulting in privacy violations and potential legal consequences;
2. **Blockchain Vulnerabilities:** The underlying blockchain technology may be susceptible to various attacks, including 51% attacks [9], and consensus vulnerabilities, which could affect the integrity and security of the shared data;
3. **Identity Theft:** Attackers may attempt to steal user SSI credentials or private keys, allowing them to impersonate users and gain unauthorized access to their verified data;
4. **Phishing Attacks:** Social engineering tactics can be used to deceive users into revealing their SSI credentials, compromising their identity and data security;
5. **Regulatory Non-Compliance:** Failing to adhere to relevant regulations (e.g., GDPR, HIPAA) when handling regulated data can result in severe legal and financial consequences;
6. **Denial-of-Service (DoS) Attacks:** Disrupting the availability of the decentralized cloud-based platform through DoS attacks can impact data sharing and system functionality;
7. **Quantum Computing Threats:** Theoretical advances in quantum computing could potentially render traditional cryptographic methods obsolete, posing a risk to data confidentiality;
8. **Credential Revocation Challenges:** Managing credential revocation effectively is crucial. Failure to revoke or update credentials promptly can lead to data exposure and misuse;
9. **Cloud Security:** The cloud infrastructure itself must be adequately secured to prevent data breaches, data loss, or unauthorized access;
10. **Interoperability Challenges:** Ensuring secure interactions with other distributed networks, platforms, or cloud services requires careful implementation to prevent security gaps;
11. **Malicious Validators:** In decentralized systems, malicious validators can approve invalid transactions or credentials, undermining data integrity and trust;
12. **Smart Contract Data Leakage:** Sensitive data may unintentionally leak through smart contracts if not designed, audited, and tested rigorously.

Fides is built with this threat landscape in mind. The design of Fides recognises potential blockchain attack vectors; their presence was one of the driving factors in implementing DRILL and DnAMS components of Fides (see section 3.1.1.1). In particular the threats #1 (Data Privacy Violations), #2 (Blockchain Vulnerabilities), #6 (DoS Attacks), #8 (Credential Revocation Challenge), #9 (Cloud security) and #12 (Smart Contract Data Leakage) from the above list are mitigated by the design of Fides which is fully decentralized and does not rely on blockchain features which are affected by the mentioned threats. On top of this the threat #11 (Malicious Validators) even if not fully tackled by Fides itself can be mitigated with the help of a regular external audit procedure, and Fides provides necessary devices to facilitate that external audit – Fides component Regulator Portal (see section 3.1.1.1) and Fides functionality “Immutable audit logs” (see section 3.1.1.2) together create a framework for semi-automated or fully automated external audit.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	35 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

The threats #3 (Identity Theft) and #4 (Phishing Attacks) from the list above does not relate to Fides itself explicitly, they may influence Fides implicitly though if an SSI component is compromised. Therefore, it is essential to ensure proper mitigation of these risks while implementing the SSI component (see section 4.1). Furthermore, implementation of continuous behavioural authentication (see section 4.3) as a part of SSI would increase the protection from these risks.

Fides is aiming to be fully legally compliant and thus aiming to mitigate also the threat #5 (Regulatory non-compliance). Further strengthening of legal compliance will be achieved by implementing TANGO task T1.4 (Data Management, Research Ethics and Legal Compliance), in particular implementing a thorough legal compliance assessment of Fides as an integral part of the task T1.4.

The threat #7 (Quantum Computing Threats) can be easily mitigated by using proper encryption algorithms which are quantum resistant. Fides itself does not impose any specific encryption algorithm, so the mitigation of this risk is as trivial as selection of a proper encryption algorithm to be used. Still, attention should be paid to choosing an implementation of the encryption algorithm – the implementation (a software library) should be chosen carefully from a list of well-accepted and properly tested implementations.

Finally, this leaves only threat #10 (Interoperability Challenges) unanswered. This is an inherent risk of heterogenous systems, which cannot be avoided by an internal means of any of the constituting components, instead this should be dealt with on an integration level – proper planning and execution of integration testing is required to mitigate it (see TANGO task T6.2 – Functional Testing and Monitoring).

3.1.4 Functional & Non-functional Requirements

Table 4: Blockchain-based Data Storage and Sharing Component Functional Requirements (Author: NOR)

ID	Description	Priority
FR-DATA-BDSS-001	The TANGO system must be able to collect and manage customer consent for every act of customer data sharing. TANGO system must be able to keep the collected consent, enforce consented policies and demonstrate it to the auditing authority.	H
FR-DATA-BDSS-002	The TANGO system must ensure that the data sharing mechanism is fully compliant with the relevant regulations concerning private data handling.	H
FR-DATA-BDSS-003	The TANGO system must ensure that the data access logs are stored immutably, in a safe and legally compliant manner, and are available to auditing authorities upon request.	H

Table 5: Blockchain-based Data Storage and Sharing Component Non-Functional Requirements (Author: NOR)

ID	Description	Priority
NFR-PER-BDSS-001	The TANGO system must deliver reasonable performance.	H
NFR-REL-BDSS-001	The TANGO system must guarantee 100% data delivery when data sharing component is involved.	H
NFR-SEC-BDSS-001	The TANGO system must guarantee that user data is being handled securely.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	36 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

3.2 Trustworthy Data Sharing [T3.2]

3.2.1 Description & Functionality

Trustworthy Data Sharing can be considered as a set of tools supporting access and usage control of data. It does not enforce usage control on its own, but it contributes to a controlled usage of data across entities.

Therefore, an additional engine for policy driven data access and usage should enforce policies to data and continuously control how data may be processed, stored, aggregated, or forwarded. Such an engine may consider technologies like Attribute Based Access Control or extensions like Usage Control. Other enforcement strategies could be a policy driven combination of different Privacy Enhancing technologies that range from anonymization, over the generation of synthetic data to differentially private mechanisms or encryption. Nonetheless, the policy engine as well as the privacy enhancing technologies are beyond the scope of T3.2. Within T3.2 three major components will be developed:

- ▶ A Data Usage Control component providing a privacy risk score.
- ▶ A Trustworthiness model for calculating a Trustworthiness Score.
- ▶ A privacy preserving recommendation module using Ubiquitous Personal Context Vectors (UPCVs).

The **Usage Control** component will calculate a privacy risk analysis score for the type of data to be shared or for operations, e.g., aggregation, joining, etc., to be done on sensitive data. based on simulated attacks/trustworthy score.

The **Trustworthiness Scoring** module shall evaluate the Trustworthiness of data sharing partners (consumer and provider). It is based on the IIC trustworthiness model utilizing system characteristics, which are evaluated as weighted sum of attributes' properties. These represent the degree of fulfilment of criteria representing policies. These characteristics may also be aggregated to a single Trustworthiness score in the same way as a weighted sum.

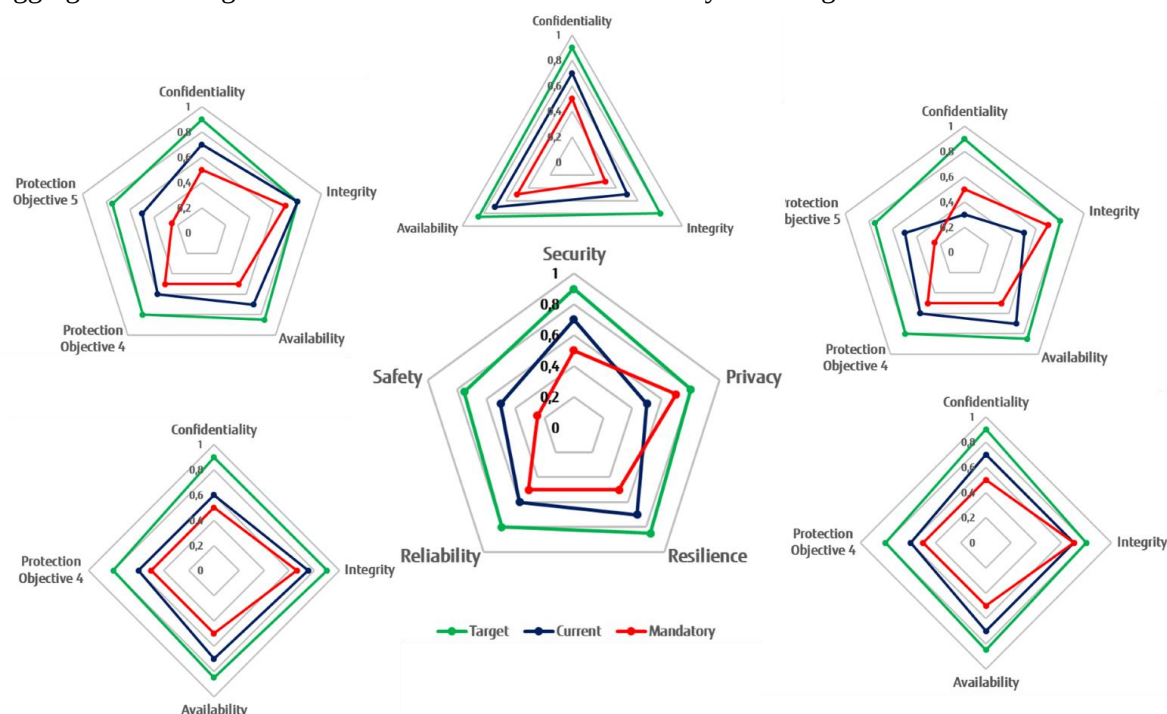


Figure 6: Trustworthiness Scoring – Staggered Schema

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	37 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

The scores shall be evaluated periodically or on demand, e.g., in case of requesting data or access to systems. This way fulfilment of Trust related criteria can be checked.

The **privacy preserving UPCV (Ubiquitous Personal Context Vector)** technology is designed for personalized recommendations, service personalization and advertising, based on user behaviour. Recommenders, in general, aim at predicting user behaviour, therefore the technology can be assumed to apply for business intelligence as well, e.g., for defining user groups and item clusters. UPCV can be implemented into a decentralized or distributed architecture, enabling information sharing between different vendors without breaching users' privacy. VTT is prepared to make a simplified recommendation service (with UI) for end-users, in case the existing platforms in piloting do not support such functionality.

Goal of the task: Provide technologies enabling trustworthy data sharing by supporting the enforcement of policies.

- ▶ **Usage Control** component: Provide a privacy risk score.
- ▶ **Trustworthiness Scoring** module: Provide a configuration of weighted characteristics, attributes and properties depending on use cases and calculate Trustworthiness scores per system characteristics according to configuration.
- ▶ **UPCVs**: Provides recommendations for end-users, such as offerings in a hotel (e.g. happy hours or recreational or cultural activities). For retail industry, it is likely to provide business intelligence, such as item clustering (including people-who-bought-this-also-bought-that) and defining different user profiles.

3.2.1.1 Internal Architecture

The Usage Control component and the Trustworthiness Scoring module rely on a preparation phase for configuration. The specific privacy risks and data usage need to be assessed for defining a privacy risk score related to the data usage. For the Trustworthiness Scoring module, the system characteristics corresponding to the use cases' applications, and their attributes need to be specified in collaboration with the use cases. In both cases this includes the specification of how the properties of the defined attributes will be evaluated on the use case side.

During operation of the **Trustworthiness Scoring** module two steps are requested.

- ▶ 1st the specific configuration of the module shall be assigned. This includes the selected characteristics, attributes, and related weightings. The attribute evaluation environment needs to be specified. Probably specific modules evaluating the attributes shall be integrated at the sharing partners.
- ▶ 2nd the attributes' evaluations shall be gathered either periodically or upon request for scoring.

The figures below illustrate this two-step process.

Configuration: The configuration process in principle shall look like the figure below:

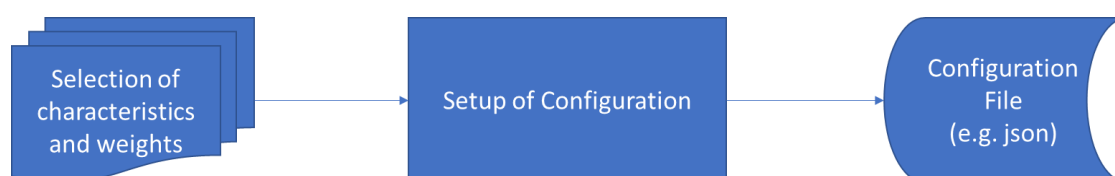


Figure 7: Trustworthiness Scoring – Configuration Process

The configuration sets the basis for the calculation of the trustworthiness score. As illustrated in Figure 6 the calculation of the Trustworthiness score follows as staggered scheme. The attributes are evaluated giving a degree of fulfilment (values between 0 and 1) and aggregated

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	38 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

in a weighted sum (sum of weights equals one) per characteristics. Overall, this may be illustrated by the following set of formulas:

$$\begin{aligned}
 & Trustworthiness(IoT_System) = \\
 & \begin{cases} 0, \text{one or more } TW(i) \text{ is below the according mandatory threshold} \\ \sum_{i \in \{Security, Privacy, Resilience, Reliability, Safety\}} W_{SC}(i) * TW(i), \text{all mandatory thresholds are met} \end{cases} \\
 & \text{with} \\
 & W_{SC}(i), TW(i) \in [0,1], i \in \{Security, Privacy, Resilience, Reliability, Safety\}, \\
 & \sum_{i \in \{Security, Privacy, Resilience, Reliability, Safety\}} W_{SC}(i) = 1
 \end{aligned}$$

The level per system characteristic (SC) may be considered as an analogous weighted combination of the degree of fulfilment of protection profiles:

$$\begin{aligned}
 & TW(i) = \\
 & \begin{cases} 0, \text{one or more } PO_{SC}(j) \text{ is below the mandatory according threshold} \\ \sum_{j=1, \dots, n_{SC}} W_{SC}(j) * PO_{SC}(j), \text{all mandatory protection objectives are met} \end{cases} \\
 & \text{with} \\
 & W_{SC}(j), PO_{SC}(j) \in [0,1], j = 1, \dots, n_{SC}, SC \in \\
 & \{Security, Privacy, Resilience, Reliability, Safety\}, \\
 & \sum_{j=1, \dots, n_{SC}} W_{SC}(j) = 1
 \end{aligned}$$

The formulas illustrate that if a mandatory characteristic or a mandatory attribute is not fulfilled, that this cannot be compensated. The two vectors of weights and evaluated fulfilment characterise the IoT system and its compliance requirements like a fingerprint. Hence, this is a kind of biometric data of an IoT system and should be protected like biometric data e.g., in communications the least necessary information should be transmitted.

The characteristics then shall be aggregated to a Trustworthiness Score in a weighted sum. Figure 8 illustrates potentially contributing modules of TANGO.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	39 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final



Figure 8: Trustworthiness Scoring – Calculation Components

Moreover, in TANGO T3.2 the **UPCV** technology shall be extended to a distributed architecture. In this distributed architecture of the UPCVs, each user and each vendor have an instance of a User Behavior Exchange Module (UBEM), which is a symmetric component - similar on user and vendor (service) side. UBEM aggregates user behavior into Volatile Random Numbers (VRN) maintaining no association of personal action history. Each instance of UBEM must have a data repository into which VRNs are collected – typically a few kilobytes total per user and per item.

In Figure 9 on the left, the user indicates an interest to an Item (*), after which vendor's UBEM and user's UBEM exchange VRNs between the respective VRN collections.

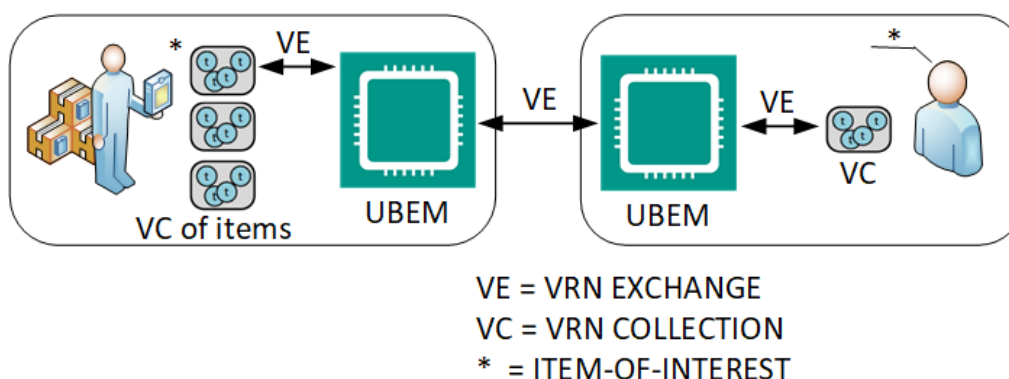


Figure 9: User interest and exchange of Volatile Random Numbers by the UBEM

In Figure 10 on the right, the user is provided by item VRN collections, among which the user selects the collection(s) that has the best match and receives item identifier(s) of the respective recommendation(s).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	40 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

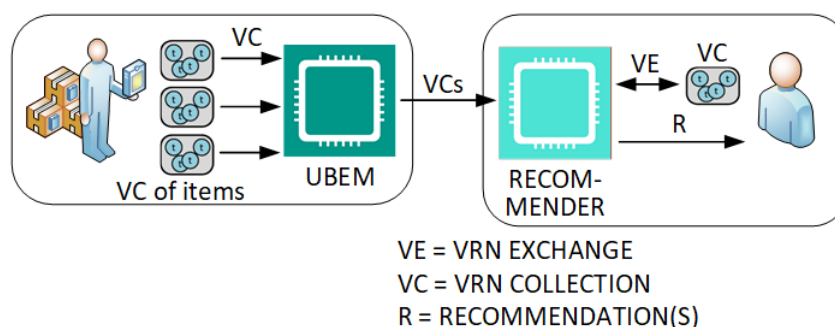


Figure 10: User Recommendations

It should be noted that UBE M itself may act as the recommender; creating recommendations may take place on either side - or by a third party. For UX, item details should then be subsequently provided to the user, based on the item identifier(s). UI/UX is use case specific.

The Figure 11 illustrates different vendors who have same or similar items to synchronize the respective VRN collections.

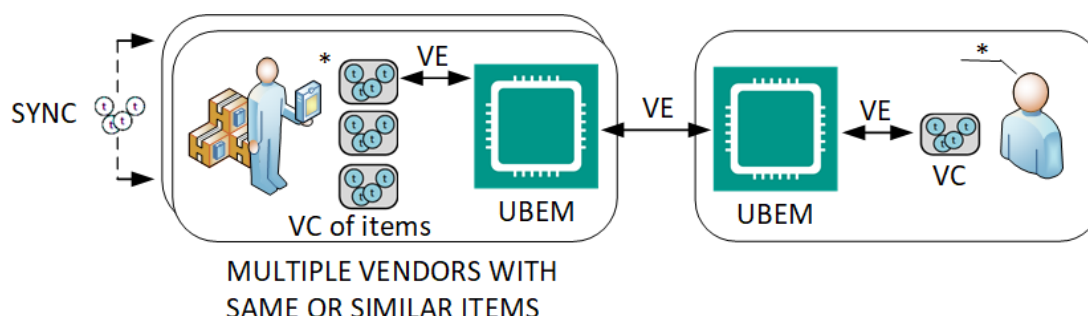


Figure 11: Synchronizing VRNs across vendors

3.2.1.2 Inputs, Outputs and Interactions with Other Components

The **Usage Control** module provides a Privacy risk score (value between 0 and 1) as output for data sharing or specific data operations. The specific inputs need to be investigated per use case.

The **Trustworthiness Scoring** Module relies on several inputs. In the configuration phase characteristics, attributes and related weights and mandatory threshold shall be specified in collaboration with the use cases. The scores per characteristics and the overall Trustworthiness score shall be calculated, aggregating the degrees of fulfilment per attribute by weighted sums using the weightings and thresholds defined during the configuration.

The **UPCV** recommendation algorithm is using user actions as its input for learning, conceptually in a format of UserID/ItemID. In a distributed implementation, each user has its own instance of the UBE M storing their VRN collection and each vendor has an instance of UBE M storing one VRN collection per item.

A) In the distributed implementation UBE Ms on user and vendor side exchange VRNs, learning is triggered by user's indicated interest in the particular item, requesting at least one VRN from the vendor, from ItemID's collection:

USER->VENDOR: {ItemID}

VENDOR->USER: {VRN}

followed by the user returning one (or more) VRN(s) from their collection:

USER->VENDOR: {ItemID, VRN}

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment	Page:	41 of 237
Reference:	D2.3	Dissemination:	PU
	Version:	2.0	Status:
			Final

For recommendations, the aim is to find the items from vendors offering VRN collections that most closely match the users VRN collection. At its simplest, the user may disclose their VRN collection to the vendor, and the vendor returns ItemIDs that have the closest matches, with similarity computations for each ItemID:

```
USER->VENDOR: {VRN collection}
VENDOR->USER: {{{ItemID, similarity}}}
```

Brackets [] indicating an array.

However, for uncompromised privacy, the recommendations may be computed solely on user's UBEM. In this case, the vendor provides VRN collections of its items for the user, ItemIDs preferably obfuscated e.g. using HMAC:

```
VENDOR->USER: {[VRN collection, HMAC(ItemID)]}
```

after which, the user requests the items that have highest similarity in their VRN collections:

```
USER->VENDOR: {[HMAC(ItemID)]}
```

and the vendor returning the respective item identifiers, in a short list

```
VENDOR->USER: {[ItemID]}
```

It should be noted that the size of a VRN collection is typically only few kilobytes per user or per item.

B) When applying UPCV for business intelligence, the implementation may be centralized, in which case the input of the recommender is UserID/ItemID pairs, preferably in randomized order. In the absence of actual user identifiers, a shopping basket identifier can be used.

```
POS->RECOMMENDER: {[UserID, ItemID]}
```

POS being an acronym of an exemplary Point-of-Sales system.

When requesting similar items, only an ItemID must be defined, the recommender returning list of similar items with similarity computations for each:

```
POS->RECOMMENDER: {RequestItems, ItemID}
RECOMMENDER->POS: {{{ItemID, similarity}}}
```

And, if any user identifiers are available, similarities between users can be computed for subsequent clustering, using as criteria a minimum similarity to control the length of the response array:

```
POS->RECOMMENDER: {RequestUsers, UserID, minimum similarity}
RECOMMENDER->POS: {{{UserID, similarity}}}
```

Or to find sizes of potential user group for a certain item:

```
POS->RECOMMENDER: {RequestUsers, ItemID, minimum similarity}
RECOMMENDER->POS: {{{UserID, similarity}}}
```

From which a size of the user group can be computed. After finding user groups of interest, items recommended for them can also be requested by presenting a request for a typical user:

```
POS->RECOMMENDER: {RequestItems, UserID, minimum similarity}
RECOMMENDER->POS: {{{ItemID, similarity}}}
```

Up to now it is not possible to provide a definitive list for the interactions of the trustworthy data sharing component with other components, but the following interactions are foreseen.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	42 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

For instance, there probably be an internal interaction of the Usage Control component providing privacy enhancing technologies and the Trustworthiness Scoring module reusing the Privacy scoring.

The Trustworthiness Scoring module evaluates properties of the data sharing partners. It thus can be used for access or usage control of systems or data. Therefore, its interactions with several components of TANGO shall be assessed during TANGO. Among those there may be:

- ▶ Usage of results of T3.3 for evaluation of PRIVACY criteria
- ▶ Usage of results of T4.1 for evaluation of criteria provided by the SSI
- ▶ Contribution to T4.3 with Trustworthiness scores
- ▶ Contribution to T4.4 with Trustworthiness scores

The UPCVs shall interact with data storage components for the VRN collections:

- ▶ size of few kilobytes per user and per item

In addition, per case (retail, smart hospitality) there are further interactions with the TANGO or the local infrastructure:

- ▶ user interests to items, i.e., user-item transactions, for recommender training (e.g., anonymous shopping baskets, browsing item presentations at hotel, actual selections)
- ▶ UBEI implementation may include libraries for Android, iOS, Windows and Linux, as well as a server implementation, to be used over an API (e.g., REST), depending on the final applications.
- ▶ if available, a user interface for presenting item selections to smart hospitality users and respective APIs

3.2.1.3 Significance and Added Value for TANGO

The technology potential of the Usage Control component needs further investigation. Nonetheless it may contribute to the Trustworthiness evaluation and an overall usage policy enforcement.

The Trustworthiness Model and module is inspired by discussions on secure cross-company communication in the German Platform Industrie 4.0 and in the Industrial Internet Consortium. There it gained high visibility and significance. The value of this component for TANGO shall be the implementation and application in the industrial use cases and extending the scope and applicability to other use cases beyond manufacturing. This way, an integrated approach on Trustworthiness shall be promoted by TANGO.

The potential of the UPCVs is outlined as follows:

- ▶ Vendors may co-operate by sharing item specific VRN collections, without compromising users' privacy.
 - This is actually a parallel with Amazon platform recommendations (people who bought this also bought that), in which a single player (i.e., Amazon itself) provides and monetizes recommendations based on user behaviour on their single platform.
- ▶ Recommendations for users are created in a multi-vendor environment while respecting privacy, by exchanging data that is not private by nature.

Last but not least, VRN collections is likely to be useful for privacy-respecting business intelligence, subject to be studied in the project.

3.2.2 Associated Pilot Use Cases

Table 6 and Table 7 provide the information of the trustworthiness and UPCV modules of T3.2 respectively with the TANGO use cases.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	43 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Table 6: Association of Pilot Use Cases with trustworthiness scoring module of T3.2

Pilot	Association
Autonomous Vehicles	The Trustworthiness Scoring is likely to provide a score for securing the connection of the vehicle to the cloud infrastructure.
Smart Manufacturing (RIAS)	The Trustworthiness Scoring is likely to provide a score for securing the connection to the shopfloor

Table 7: Association of Pilot Use Cases with UPCV module of T3.2

Pilot	Association
Smart hospitality	UPCV provides recommendations for end-users, such as offerings in a hotel (e.g., happy hours or recreational or cultural activities).
Retail	For retail industry, UPCV is assumed TO be applicable for business intelligence, such as item clustering (including people-who-bought-this-also-bought-that) and defining different user profiles.

3.2.3 Threat Landscape

Threats shall be assessed during further development of the related modules. Typically, threats will result from the used operations environment, e.g., containers, OS, libraries, etc. over time caused by zero-day exploits or weak information security procedures, e.g., related to updating or patching of critical vulnerabilities.

Moreover, secure communication is highly important avoiding poisoning of data during transmission. Considering especially the UPCVs disclosing too many VRNs in one session may compromise anonymity of a returning customer, e.g. “this anonymous customer A was presumably the same as the anonymous customer B last week”.

The major prevention and tackling mechanisms shall be related to the used components and their vulnerabilities. Hence, a common vulnerability management shall be considered starting with the design phase.

Moreover, related to the UPCV minimizing the number of VRNs disclosed in a session shall prevent the disclosure of sensitive information. The latest version of the algorithm has been proved effective even when only a single VRN has been disclosed.

- ▶ This is to be on the safest side; disclosing more VRNs may require further measures for uncompromised privacy.
- ▶ When a user is known by the vendor, however, the number of disclosed VRNs does not have any adverse effect on privacy, at least not within reasonable limits and visiting frequency. (Can be further studied during the course of the project.)

3.2.4 Functional & Non-functional Requirements

The functional requirements for this component are presented in Table 8. This list reflects the link of each of the modules with the use cases: **Usage Control** is linked to many use cases, the **Trustworthiness Scoring Module** is linked to use cases Autonomous Vehicles and Smart Manufacturing, and the **UPCV** module is linked to the use cases Smart Hospitality and Retail.

Table 8: Trustworthy Data Sharing Component Functional Requirements (Author: FHG, FUJ_DE, VTT)

ID	Description	Priority
FR-DATA-TDS-001	Simulated privacy attacks based on predefined adversary settings against the data sharing process or the shared data to estimate the risk of possible privacy leakages. The privacy risk	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	44 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

	analysis must be performed to aid in the creation of a suitable usage control policy that is then enforced later.	
FR-DATA-TDS-002	The TANGO system must provide a module configuring the Trustworthiness calculation and a module calculating the Trustworthiness scores based on this configuration for the vehicle and the CAVride system.	H
FR-DATA-TDS-003	Thus, the TANGO system must provide a module configuring the Trustworthiness calculation and a module calculating the Trustworthiness scores based on this configuration for the Data Gate Keeper system.	H
FR-DATA-TDS-004	The TANGO system should provide a data repository of few kilobytes for each end guest and each vendor (i.e. hotel, activity provider etc.).	M
FR-DATA-TDS-005	The TANGO system should provide user transaction data, e.g., shopping baskets.	M

Table 9: Trustworthy Data Sharing Component Non-Functional Requirements (Author: FHG, FUJ_DE, VTT)

ID	Description	Priority
NFR-SEC-TDS-001	The TANGO system could provide the continuous Privacy Risk Score to a monitoring system, which could visualize the Privacy for easy evaluation.	L
NFR-SEC-TDS-002	The TANGO system could provide the continuous Trustworthiness Score to a monitoring system, which could visualize the Trustworthiness of the vehicle or the platform for easy evaluation by a user.	L
NFR-SEC-TDS-003	The TANGO system could facilitate the assessment of the regular operations of the shopfloor	L
NFR-UI-TDS-001	The TANGO system should provide hotel guests recommendations in a user-friendly personalized order	M
NFR-UI-TDS-002	The TANGO system should provide business intelligence from retail locations that are not per se allowed to share customer data with each other.	M

3.3 Confidentiality and Privacy by Design [T3.3]

3.3.1 Description & Functionality

The confidentiality and authentication of shared data will be supported by a module for creation and application of sticky policies and signatures (CP module). Data will be encrypted according to an attribute-based policy set by the data owner. Potential consumers will need to retrieve keys associated to their identity attributes and will only be able to decrypt data if their attributes (i.e., the respective keys) fulfil the sticky policy. Additionally, data can be signed, e.g., using the SSI information as a basis will be explored, for verification of integrity and provenance of the data. The cryptographic operations can be supported by the ABE Toolset, which integrates with open-source state-of-the-art implementations of functional encryption schemes:

- 1) <https://github.com/fentec-project/gofe>
- 2) <https://github.com/fentec-project/CiFEr>

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	45 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

The main data sharing mechanisms that CP module will include are:

- ▶ Cryptographic tools (encryption/decryption/signatures/verification)
- ▶ Access policy enforcement (in particular, access to the plain shared data, as encrypted data retrieval will be done at the cloud storage, which is a generic TANGO functionality and not CP itself).

The way sticky policies will be instantiated is with CP-ABE, consequently, from now on in TANGO, CP-ABE and CP module are synonymous and imply sticky policies.

Policy-based access control will be used to address user consent (UC). Attribute-based policies will be used to enable fine-granular checks based on identity attributes (e.g., coming from project's SSI solution). As an additional protection measure, policies may require the application of PET (e.g., Anonymization etc.) to data.

The main features to tackle UC will be:

- Policy to describe the info to be shared by the entity (T3.2/T3.3, T4.1... T8.6). (UC)
- User consent management (linked to access policies). (UC)

3.3.1.1 Internal Architecture

The UC is not a module (i.e., single package that works as a standalone and receives input/provides output) per se, and instead it is an approach based on the definition of attribute-based policies to address the authorization needs. Nonetheless, tools for policy visualization and parsing (human<->machine readable) will be needed. It may depend heavily on specifics of the pilot (how is consent actually gathered, who is using the system--actual data owner, data controller/processor etc.)

The CP module will have two separate subcomponents:

- ▶ Client-side component: submodules for encryption, decryption, policy engine. It interfaces with SSI and secure key management for achieving its functionalities.
- ▶ The server-side component is mainly an issuance module that interfaces with SSI for checking validity of issuance requests. The ABE Toolset can be the main asset for cryptographic operations in the corresponding modules.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	46 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

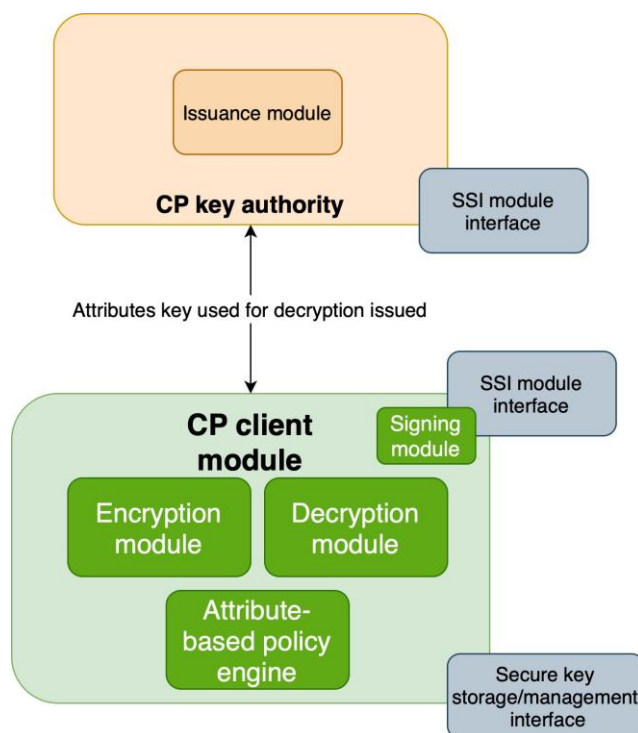


Figure 12: CP module architecture

CP-ABE will follow the beside process flow:

- ▶ **SSI initial identity:** following SSI process defined in section 4.1.1., data consumer will need to get a verifiable credential (VC) by proving its identity to the SSI issuer.
- ▶ **Data publishing:** using its public key, data provider will encrypt its data together with its sticky policies so that only desired users can have access to it. This encryption may also be signed with its signature to be investigated making use of SSI keys. Besides the encrypted data, some metadata regarding the encryption have to be published on the cloud storage infrastructure.
- ▶ **Data retrieval:** in order to access the encrypted data, data consumer must obtain the access policy from the cloud storage (possibly together with some other enforcement mechanism, e.g., IDSA expiration date) and responds with a Verifiable Presentation (VP will benefit from ZK (Zero Knowledge) for added privacy) according to it. This process will give the data consumer access to the encrypted data provided previously by the data producer.

In order to decrypt data, data consumer will need to interact with CP-ABE KA. The latter will provide CP-ABE keys needed for decryption if the data consumer fulfils attribute prerequisites (by means again of VP).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	47 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

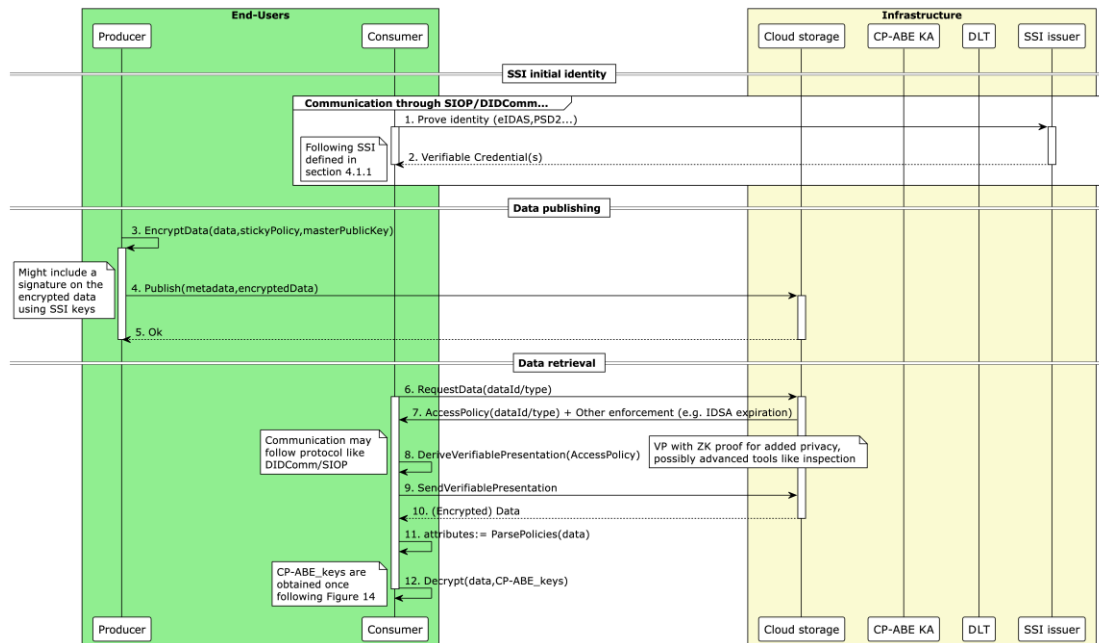


Figure 13: CP-ABE workflow

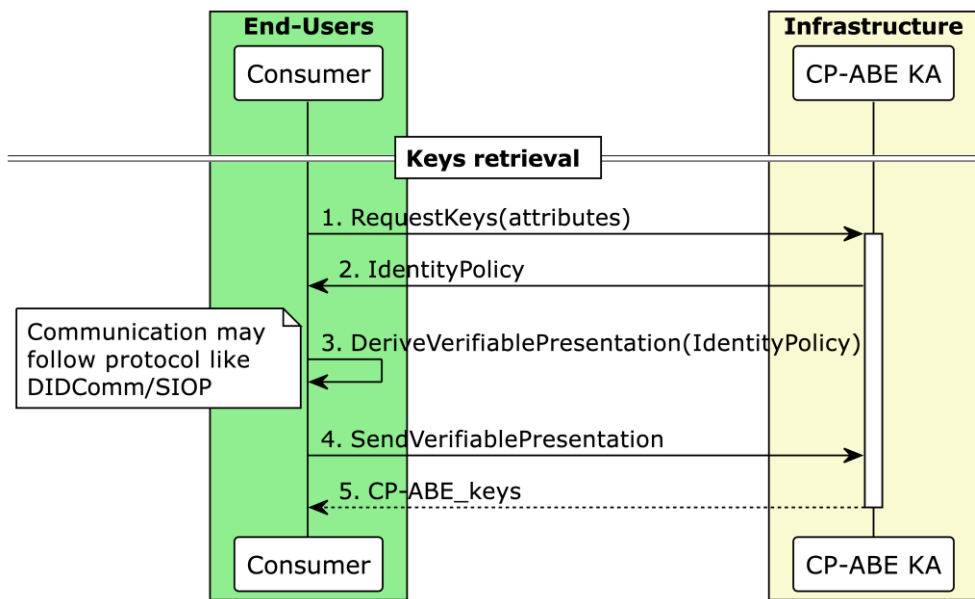


Figure 14: CP-ABE Keys Retrieval

3.3.1.2 Inputs, Outputs and Interactions with Other Components

The CP component will take as input both data in plain, user/consent input (i.e., description of what policy should be applied to the data) as well as verification results from SSI module (verification interaction for key issuance). It will provide as output encrypted and signed data together with its sticky policies, which will be shared in the data sharing tool. Besides, this data will be decrypted and authenticated for the consumer.

User consent will be reached by using as inputs user consent forms and attribute-based policy of some request. The result will be machine-readable attribute-based policies associated to user

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment	Page:	48 of 237
Reference:	D2.3	Dissemination:	PU
	Version:	2.0	Status: Final

consent and their respective visualizations. It is expected that the CP and UC components will have the following interactions with other TANGO components.

CP:

- ▶ SSI (T4.1) for verification of identity attributes needed for issuance of attribute keys for decryption. Another potential interaction with SSI will be its use for signing (and verifying) the data for authentication.
- ▶ Data sharing tasks (e.g., T3.2) may be useful for full access control, e.g., IDSA proposes an expiration date for shared data.

UC:

- ▶ SSI (T4.1) – Attribute-Based policies will be used for authorization through SSI tools in the project.
- ▶ Policies/user-consent will also be necessary in the processes related to other tasks like T3.2.

3.3.1.3 Significance and Added Value for TANGO

The CP module may be used in use cases where extra data confidentiality mechanisms are needed or desirable, allowing data owners to establish policies that will be enforced throughout the data lifecycle in the data sharing platform.

The UC module will enable adherence to GDPR by allowing the application of user consent directives through fine-grained access control mechanisms.

3.3.2 Associated Pilot Use Cases

In the following, the relation of the confidentiality and privacy by design functionalities with the different TANGO use cases is highlighted.

Table 10: Association of Pilot Use Cases with confidentiality and privacy by design functionalities of T3.3

Pilot	Association
Smart Hospitality	Attribute-based policies can be implemented to safeguard data shared across CRM boundaries, ensuring controlled access throughout its lifecycle. This allows the user or data controller to have granular control over determining who can access the data at various stages.
Autonomous Vehicles	To maintain privacy while sharing sensitive data, such as user preferences or maps, proper access control and user consent mechanisms are implemented to ensure compliance with GDPR rights. Sticky policies and related identity-based techniques are employed to enable decryption of data only by authorized entities. For example, only specific cars or authorized clients are granted decryption capabilities for accessing user service preferences or maps.
Smart Manufacturing (FMAKE)	The system supports access control for data, specifically printing job information tailored to specific clients. Attribute-based and sticky policies are utilized to establish precise control over who can access the encrypted data based on their identity. For instance, all data from a printing job can be uploaded to a sharing platform, but each component is encrypted with specific policies, ensuring that only authorized users can decrypt and access the relevant data, such as data requested by a specific client.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	49 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Smart Manufacturing (RIAS)	<p>The system provides support for access control to infrastructure and equipment by enabling the definition of attribute-based policies.</p> <p>This allows for the establishment of policies that utilize identity data in the authorization mechanism, determining which digital resources should be accessed by specific individuals, particularly in cases of remote access from vendors or employees. This ensures a secure and controlled environment for granting access to authorized personnel based on their defined attributes.</p>
Public Administration (VISAR)	<p>The system ensures adherence to privacy principles in data procedures, with a specific focus on GDPR compliance. This is achieved through robust data access control mechanisms and user consent mechanisms. Sticky policies and related identity-based techniques are employed to enforce restrictions on data retrieval, in addition to database access control.</p> <p>These measures enable fine-grained control, ensuring that only authorized actors, such as the corresponding manager and team leaders, can access data from a specific case. This safeguards user data and maintains privacy while adhering to regulatory requirements.</p>
Retail	<p>The system guarantees the confidentiality of data and empowers partners with absolute control over how it is shared.</p> <p>By utilizing CP-ABE and related identity-based techniques, the system ensures that only authorized entities can decrypt the data. This provides an additional layer of protection for data at rest, in addition to access control measures on cloud sharing platforms.</p> <p>Partners retain complete control over their shared data by establishing policies governing who can access it. This ensures that data shared through the platform, such as analytical data from specific physical stores, remains safeguarded according to the partners' preferences and requirements.</p>

3.3.3 Threat Landscape

Concerning the CP module, the following threats are considered.

- ▶ Denial of service of Key issuance.
- ▶ Theft of attribute keys from end-user device.
- ▶ Issuance of attribute keys to inappropriate actors.

To tackle the aforementioned threats, the following mechanisms are considered.

- ▶ Protection of key issuance servers, e.g., through firewalls or throttle, and potentially replication of servers.
- ▶ The CP-ABE keys, along with other sensitive material in the project, will be securely stored in the device thanks to TANGO tools (e.g., T3.4).
- ▶ The issuance of keys will be protected through an authorization process that relies on the project's identity management framework, in particular through verification of verifiable presentations as supported by T4.1.

3.3.4 Functional & Non-functional Requirements

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	50 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Table 11: Confidentiality and Privacy by design Component Functional Requirements
(Author: UMU)

ID	Description	Priority
FR-DATA-CPD-001	The TANGO system must allow an administrator or policy authors to define the policies based on attributes.	H
FR-DATA-CPD-002	The TANGO system must enforce access control decisions based on policies evaluation.	H
FR-DATA-CPD-003	The TANGO system must apply fine-grained control over data access and usage.	H
FR-DATA-CPD-004	The TANGO system should define the appropriate security controls at each stage of the data lifecycle.	M
FR-DATA-CPD-005	The TANGO system must involve mechanisms to enforce policies and restrictions.	H
FR-DATA-CPD-006	The TANGO system must support the enforcement of defined policies by an encryption/decryption mechanism.	H

Table 12: Confidentiality and Privacy by Design Non-Functional Requirements (Author: UMU)

ID	Description	Priority
NFR-SEC-CPD-001	The TANGO system must prioritize and enforce robust security measures to protect data.	H
NFR-SEC-CPD-002	The TANGO system must ensure the privacy of individuals and sensitive data aligned with GDPR.	H
NFR-PER-CPD-001	The TANGO system should not introduce significant performance overhead that could impact system responsiveness.	M
NFR-REL-CPD-001	The TANGO system must ensure continuous enforcement of policies.	H

3.4 Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms [T3.4]

3.4.1 Description & Functionality

Self-encryption and decryption can be used as efficient methods to protect data generated by different use cases. Multi-factor information recovery supports secure access to user and other data on cloud servers if audit services are required. Only individuals who provide and match multiple factors can access sensitive information on the cloud server and thus reducing the risk of data theft or misuse.

As for the technical component ID-based self-encryption, its main processes include the following:

- ▶ We first define the data format - which data we are going to protect and how these data are organized in what format;
- ▶ We then divide the data into multiple data chunks.
- ▶ We further select an efficient and lightweight symmetric encryption as a main encryption tool along with a collision resistant hash function;

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	51 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ We merge the data owner's with ID and link ID to the data chunks so that data users or auditor is able to know who owns the data, maintaining the ownership of the data;
- ▶ The data chunks are encrypted into an individual encryption copy.
- ▶ During the decryption, with a valid key, the data chunks can be fully recovered and the ID can be seen also.
- ▶ Once ID is revealed, we can trace who the data chunk belongs to.

As for the multi-factor data recovery mechanism, we are going to provide the following functionalities:

- ▶ Given a data decryption key, we use a secret sharing scheme to divide the key into multiple pieces.
- ▶ We define these pieces as factors. And these factors are considered as the data recovery rights.
- ▶ We reflect a single factor into a secret share and set the bar - how many factors should be presented to fully recover the key.
- ▶ With sufficient number of the factors, one can correctly recover the key. And we note that multi-factor could be obtained by one user but also shared among multiple users.

3.4.1.1 Internal Architecture

The self-encryption contains three main steps (Figure 15). The first step is the data chunk creation, in which the initial data (plaintext) is divided into identical data chunks ($C_0, \dots, C_{n-1}, C_n$). The second step contains three main phases. First, in the first phase the hash value of each data chunk needs to be calculated. The cryptographic hash value of a chunk is the unique representation or the so-called digital fingerprint of the chunk. In practice, the hash value is at least 256 bits long; for example, SHA [10], Keccak [11], or Blake2 [12] cryptographic hash function families can be applied in order to calculate the corresponding hash values. Next, the second phase of this step is the generation of the secret keys and the initialization vectors for chunks' encryption. The AES-128 block cipher is used to encrypt all data chunks, which requires a secret key (Key) of 16 bytes and an initialization vector (IV) of 16 bytes. The hash values computed in the previous phase are used to create the initialization vectors and keys. For encrypting chunk (C_n), the previous chunk's (C_{n-1}) hash is required for the AES block cipher function. The first 16 bytes of the previous chunk's (C_{n-1}) hash value serves as the key and the last 16 bytes as the initialization vector for the AES block cipher. Also, XOR obfuscation values ($X_0, \dots, X_{n-1}, X_n$) must also be computed for all chunks at this same phase. The obfuscation value corresponding to chunk (C_n) is determined by concatenating the hash of the current chunk (C_n) and the hash of the first of the two previous chunks (C_{n-2}). The third phase of the second step is the AES encryption of the chunks using the corresponding keys and initialization vectors. After obtaining the encrypted chunks, obfuscation is applied to each of the AES encrypted chunks (e.g., $\text{Enc}(C_n)$). In this last step of the self-encryption, the encrypted chunks are Xor-ed with the obfuscation values computed in the previous step in order to obtain the final encrypted chunks (e.g., EX_n). It must be noted that one obfuscation value (e.g., X_n) is 64 bytes long, contrarily to the size of one AES encrypted chunk, which can be higher than 64 bytes. If the AES encrypted chunk's size exceeds 64 bytes, the obfuscation value is rotationally padded by itself until it achieves the same length as the AES encoded chunk's size. The last stage of the self-encryption is the data map creation, which can be represented as a table. The left column of the table contains the hashes of the data chunks, which are required in order to be able to determine the keys and initialization vectors. The data chunks' hashes can be considered the secret keys for the self-encrypted data, which means that these values should not be publicly shared. The right column of the table contains the hashes of the final encrypted chunks. The final encrypted data chunks' hash values can be considered the pointer to the final

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	52 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

encrypted data chunks allowing their storage in different locations. Storing data chunks in different locations makes it harder to retrieve the totality of the chunks.

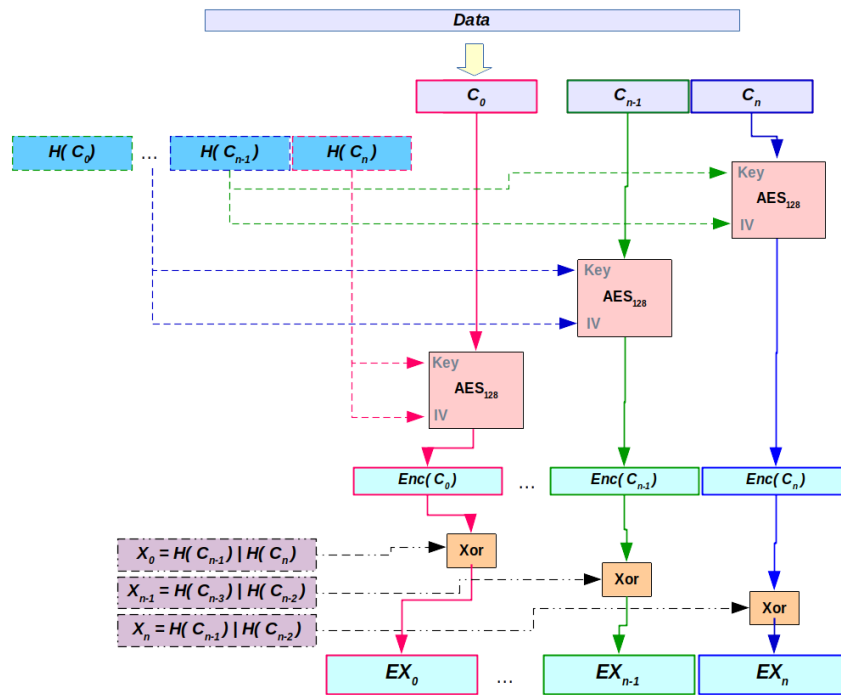


Figure 15: Self-Encryption workflow

3.4.1.2 Inputs, Outputs and Interactions with Other Components

The inputs for this component are the data, the user's identity or/and the user's private key. The output is one whole piece of the encrypted data.

The envisioned objective is to apply the above self-encryption and multi-factor recovery scheme with the following technical components. First, encrypted data should be considered stored and indexed somewhere. We may consider the decryption key could be somewhat protected on the blockchain component, e.g., Fides, related to T3.1. Also, our component could be combined with the use of policy-based data encryption and ABE in T3.3 in the sense that we may use ID-based self-encryption to encrypt the data itself, while using ABE to protect the decryption key, which is the common use of hybrid encryption, combining symmetric and asymmetric encryption techniques.

3.4.1.3 Significance and Added Value for TANGO

TANGO's self-encryption and decryption technology do not require user intervention or a password for the encryption system. This is a great convenience for our pilot use cases. Besides, the encryption here aims at creating cipher-text (encrypted) objects that are extremely strong and closer to perfect in terms of reversibility, as opposed to known encryption ciphers available today. Finally, it provides sufficient encryption security and offers good system performance, saving storage space. As for our key recovery technique using secrete sharing, this mechanism can support data recovery even when the key is lost or stolen, and this even supports third party auditing.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment	Page:	53 of 237
Reference:	D2.3	Dissemination:	PU
	Version:	2.0	Status:
			Final

3.4.2 Associated Pilot Use Cases

Table 13: Association of Pilot Use Cases with self-encryption and decryption techniques of T3.4

Pilot	Association
Autonomous Vehicles	Self-encryption and decryption can be used as efficient methods to protect data generated and collected by the autonomous vehicle's sensors and systems. This includes data such as location, speed, and sensor readings, as well as data that is transmitted to other vehicles or to the cloud server for further processing. Multi-factor information recovery supports secure access to user and autonomous vehicle data on cloud servers if audit services are required. Only authorized individuals can access sensitive information on the cloud server and thus reducing the risk of data theft or misuse.
Smart Manufacturing	In smart manufacturing, self-encryption and decryption, along with multi-factor information recovery, provide robust data protection. Production metrics, equipment status, quality control data, and supply chain information are encrypted to ensure confidentiality. Authorized individuals can decrypt and access the data securely. Multi-factor information recovery adds an extra layer of security, requiring multiple authentication factors for accessing and managing critical data stored on cloud servers. By implementing these techniques, smart manufacturing systems mitigate the risks of data breaches and unauthorized access, enabling secure collaboration and protecting the integrity of valuable manufacturing data.
Retail	Self-encryption and decryption can be used in retail sector's data sharing. For example, self-encryption and decryption can simplify the process of sharing data between retailers, suppliers, and other third-party partners by removing the need for complex encryption systems. Besides, self-encryption and decryption can enhance security by eliminating the need to store sensitive data in an unencrypted form, reducing the risk of data breaches. If audit services are required, multi-factor information recovery supports secure access to e-business data on various systems and data warehouses.

3.4.3 Threat Landscape

The following threats have been considered for the design of self-encryption and decryption technologies.

- ▶ **Key management:** To ensure that the keys are kept secure and protected from unauthorized access or theft, the keys of the self-encryption must be carefully protected and stored.
- ▶ **Attacks aiming to cryptographic primitives in the technologies:** external attackers may try to break AES, Hash function and secret sharing schemes we use.
- ▶ **Recovery factors:** How many factors are needed in multi-factor information recovery needs to be carefully considered and designed. A wrong design will lead to misuse and system vulnerabilities that affect security.
- ▶ **Potential corruption of factors:** an insider attacker (e.g., within a company being given a factor) could possibly corrupt or compromise other factors owned and possessed by other parties so as to recover the key.

Prevention mechanisms for these threats are planned to be integrated in our technology. For example, we use ID-based encryption/decryption so that no additional secret key storage scheme is generally required. We aim to produce lightweight encryption solutions, which should match the performance requirements of the TANGO platform in terms of computational

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	54 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

resources and processing data. We will carefully design the recovery factors to handle various scenarios.

3.4.4 Functional & Non-functional Requirements

Table 14: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms Functional Requirements (Author: DUT)

ID	Description	Priority
FR-DATA-SEDT-001	The TANGO system must implement robust data encryption techniques to protect sensitive information generated and collected by autonomous vehicles or smart manufacturing systems. Encryption ensures that data remains unreadable to unauthorized parties and maintains its confidentiality throughout storage and transmission.	H
FR-DATA-SEDT-002	The TANGO system should support multi-factor authentication mechanisms, requiring users to provide multiple authentication factors such as passwords, biometrics, or physical tokens to access sensitive data or system functionalities. This adds an extra layer of security, ensuring that only authorized individuals with valid credentials can gain access.	M
FR-DATA-SEDT-002	The TANGO system should implement encryption algorithms and techniques that are highly efficient in terms of computational resources and processing time. The encryption process should be optimized to minimize any performance impact on the overall system functionality.	M

Table 15: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms Non-Functional Requirements (Author: DUT)

ID	Description	Priority
NFR-MAIN-SEDT-001	The TANGO system must be designed with a modular architecture, where the self-encryption and decryption component is a self-contained and easily maintainable module. It should be capable of independent updates or enhancements without affecting the overall system functionality.	H
NFR-PER-SEDT-001	The TANGO system should exhibit high-performance efficiency in terms of encryption, decryption, and multi-factor information recovery operations. The component should be optimized to minimize processing time, computational resources, and memory usage, without compromising the overall system performance.	M
NFR-COMP-SEDT-001	The TANGO system should demonstrate interoperability with commonly used communication protocols and data formats in the autonomous vehicle or smart manufacturing domain as far as self-encryption and decryption techniques are considered. It should seamlessly exchange information with other systems to facilitate smooth integration and collaboration.	M

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	55 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

4 Distributed Trust Management Framework

4.1 Self-sovereign Identity Management [T4.1]

4.1.1 Description & Functionality

Self-Sovereign identity (SSI) Solution comprises of an end user SSI Wallet (aka Personal Credential Manager) for storing and presenting Verifiable Credentials (VCs) [12] and an SSI Agent for issuing and verifying VCs.

It is envisioned that the SSI solution will be used for TANGO. The solution offers a high level of assurance supporting de-materialisation of identity supported on international and European standards aimed at high usability, interoperability and trust based on a decentralized architecture. The highlights of the SSI Solution are as follows:

- ▶ Strong alignment with international standards & EU initiatives (W3C, OpenID, EBP, EBSI-ESSIF).
- ▶ W3C Decentralized Identifiers providing decentralized PKI for enabling trust [13].
- ▶ W3C Verifiable Credential data model with Holder, Verifier & Issuer approach [14].
- ▶ A user-centric approach which provides increased trust & transparency.
- ▶ Enhanced GDPR compliance (e.g. ZKP).
- ▶ Enabler of new business models for eTransactions based on Verifiable Credentials.
- ▶ Enabling new model of electronic transactions based on user-controlled W3C Verifiable Credentials allowing data minimisation and highly automated validation checks.

The main envisioned functionalities that will be supported are as follows:

- ▶ Issuer of Person, Organization Members and Organization entity credentials in line with recognised schemas.
- ▶ Support verification of Verifiable Credentials for Persons and Organizations.
- ▶ Support for public Decentralized Identifiers for Issuers, Organizations and IoT Devices.
- ▶ Support of Decentralized Identifier (DID) Authentication (separate IoT Device SSI component).
- ▶ Enable granular disclosure of personal data.
- ▶ Allow end-users to easily manage, send and issue verifiable credentials.

Granular disclosure includes both selective disclosure (disclosing only part of a verifiable credential through ZKP privacy preserving mechanisms) and other data minimization functionalities (e.g., binary disclosure, data fingerprinting, tracking of revealed data, minimization of data revealed by combinations of disclosures, repudiability, and automatic revocation of access rights).

The overall TANGO architecture integrates multiple GAIA-X components, several of which have specified and implemented an SSI Solution based on Hyperledger ARIES [13], which is built on top of the layer 2 DIF DIDCOMM protocol [14]. However, the latest specification of Gaia -X makes use of the OpenID Connect for Verifiable Presentations (OIDC4VP) [15] and Self-Issued OpenID Provider v2 (SIOPv2) [16], which are built on the well-known standards of OpenID. The Data Spaces convergence paper [17] also specifies compatibility with eIDAS and EBSI trust anchor framework and through the use of SIOP and OIDC4VP.

The DIDCOMM protocol is not interoperable with these OpenID standard specifications. Therefore, considering these divergent standards and requirements to integrate with GAIA-X, the exact technologies to be used have not been finalized in TANGO. Alternatives currently under active consideration include using existing software such as Hyperledger Aries,

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	56 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

supported by released GAIA-X components and also partner's assets; using the Eclipse Dataspaces Connector if the planned integration with SSI is completed early enough; and/or developing new SSI components based on OpenID standards and open-source technologies to support Verifiable Credentials. In any case, the chosen technologies should integrate with the GAIA-X infrastructure so to, for example, be able to support Service Descriptions as Verifiable Credentials. It is worth noting that GAIA-X federation services have been handed over to the Eclipse foundation and is called Eclipse-XFSC [18].

4.1.1.1 Internal Architecture

SSI Solution for persons and organization members:

The SSI Solution for supporting persons is composed of various sub-modules, depicted in Figure 13, including the core components:

► SSI Agent

The SSI Agent can be deployed as an Issuer / Verifier of Verifiable Credentials for organizations acting only as recognized issuers (e.g. of national eIDs) or as organizations that act as Service Providers who can issue their members with member credentials and also verify verifiable credentials submitted to them for authentication/authorization purposes from members and end users.

► SSI Wallet

Secure end user wallet that securely stores the user's Verifiable Credentials, provides a User Interface for the user to manage their credentials and provides the cryptographic operations when submitting & receiving Verifiable Credentials including handling of the Decentralized Identifiers.

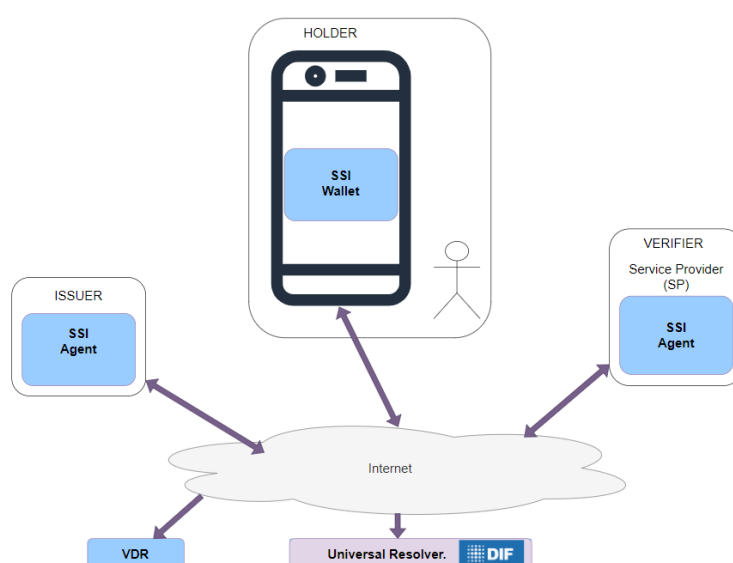


Figure 16: Illustration of the individual modules of the SSI component

IoT Device SSI component:

The IoT Device authorization flow is described below and included in Figure 17.

1. The service logic can proceed to request the device to be authorized to access a Service Provider with the device's public DID.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	57 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

2. This triggers a request from the SP to be authorised to the IoT Device SSI for the devices DID.
3. The IoT Device SSI will then send a DID Authentication request to the SP authentication endpoint published in the IoT Devices DID Doc.
4. Once authenticated the IoT Device SSI will return a signed ID Token to the SP.
5. The SP checks that the IoT Device's DID is previously onboarded against the claims received inside the ID Token and if so, creates an access token for the IoT Device.
6. The SP returns the access token to the IoT Device for establishing trusted communications.

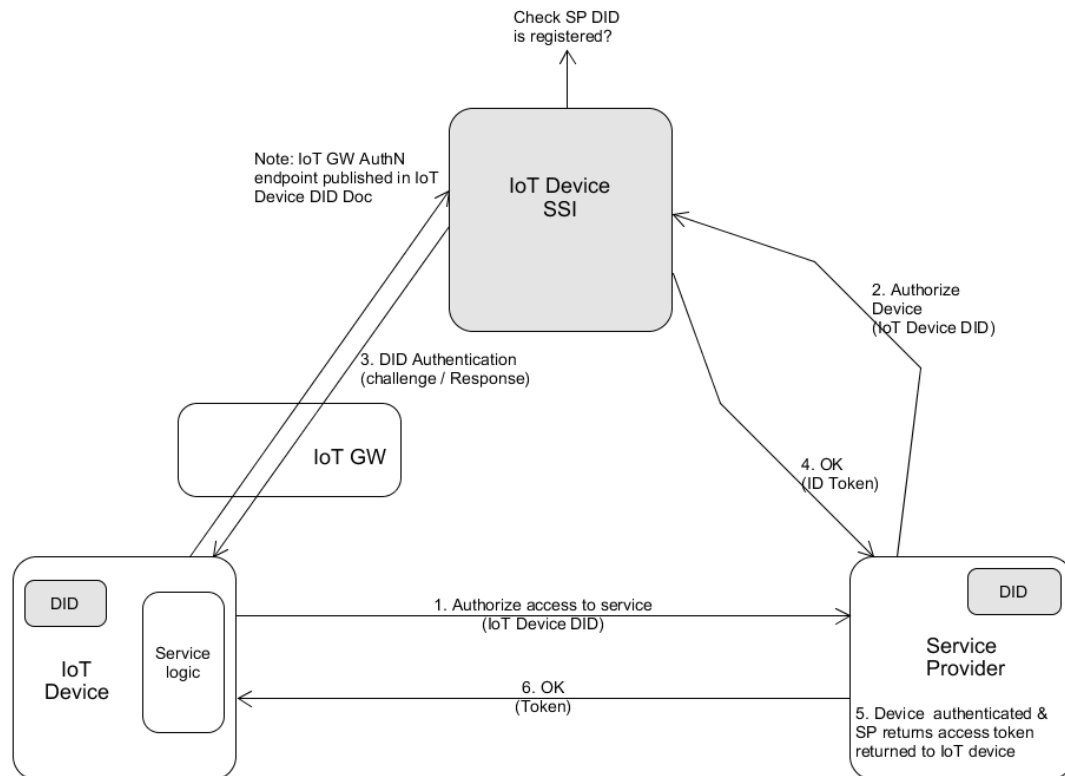


Figure 17: IoT Device SSI component and interwork

4.1.1.2 Inputs, Outputs and interactions with other components

SSI Solution for persons and organizations:

On a high level, SSI management has as its inputs:

- ▶ User information: personal information such as credentials.
- ▶ Sources of trust: actors that are able to verify the veracity of user information.
- ▶ Service provider requirements: ways for service providers to utilise verified user information.

And as outputs:

- ▶ Verifiable Credentials (VCs): user information that has been verified as true by a source of trust.
- ▶ Issuing credentials: the ability to handle and send these verifiable Verifiable Credentials.
- ▶ Proofs of credential verification: a way of checking the veracity of credentials.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	58 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

On a more concrete level, the two sub-components (SSI Agent and SSI Wallet) operate as follows.

The SSI Wallet:

- ▶ Securely handles/stores cryptographic material pertaining to its Decentralized Identifiers.
- ▶ Stores the verifiable credentials that has been issued to it.
- ▶ Replies to verifiable credential requests from a verifier SSI Agent by performing Verifiable Presentations.
- ▶ Supports ZKP privacy preserving Verifiable Credentials.

SSI Agent API operations inputs will be the following:

- ▶ Create and register a DID for the SSI Agent.
- ▶ Issue credential (Request to issue a Verifiable Credential to a Person or Organization).
- ▶ Present proof (Request to a person or organization to present a Verifiable Credential).
- ▶ Support ZKP privacy preserving Verifiable Credentials.

The following are interactions with TANGO components that are foreseen:

- ▶ ZKP Library in wallet for creating ZKP proof and broker/agent for issuing ZKP Verifiable Credential.
- ▶ Onboarding interface to issue onboarded identities as specified in T4.2.
- ▶ Attribute Based Encryption (ABE) Key Issuer from T3.3 to be able to prove identity claims to be authorized to obtain an ABE Key.
- ▶ Authentication based on OIDC / OAUTH2 that manages an ID Token.

The SSI Broker /Agent APIs output will be a webhook to return results of the requested API operations.

The inputs, outputs and interactions of the **IoT Device SSI component with other components** are described in section 4.1.1.1 within the overall view of its internal architecture and its client and server process flow.

4.1.1.3 Significance and Added Value for TANGO

These SSI tools will be available for TANGO to implement in a variety of ways. TANGO can provide them to stakeholders e.g., Service Providers, Data providers to integrate in their infrastructure or TANGO can provide a centralized SSI platform for use by the TANGO stakeholders depending on the project needs.

Added value of supporting SSI is that it is being promoted in EBSI, GAIA-X and EU DIW for authenticated access and, at the same time, it gives end users full control of their identity and how to share it, while removing dependencies on identity providers such as Google and Facebook.

4.1.2 Associated Pilot Use Cases

All the use cases in pilots use SSI in some form, at least for the purpose of identifying users. Other pilots utilise other functionalities provided by SSI as described in the following Table. Organisation service descriptor VCs will be offered for any pilots integrating with GAIA-X.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	59 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Table 16: Association of Pilot Use Cases with SSI functionalities of T4.1

Pilot	Association
Smart Hospitality	User login. User credentials. User key data. <ul style="list-style-type: none"> ▶ Issue User with eID (simulating national eID) ▶ Verify Verifiable Credentials submitted to the Organizations (e.g. Service Providers) & (from a user SSI Wallet)
Autonomous Vehicles	<ul style="list-style-type: none"> ▶ Issue User with eID (simulating national eID)
Smart Manufacturing (FMAKE)	<ul style="list-style-type: none"> ▶ Issue Service Provider staff with Organization Member credentials ▶ Verify Verifiable Credentials submitted to the Organizations (e.g. Service Providers) & (from a user SSI Wallet)
Smart Manufacturing (RIAS)	<ul style="list-style-type: none"> ▶ Issue Service Provider staff with Organization Member credentials ▶ Verify Verifiable Credentials submitted to the Organizations (e.g. Service Providers) & (from a user SSI Wallet)
Public Administration (VISAR)	<ul style="list-style-type: none"> ▶ Issue Service Provider staff with Organization Member credentials ▶ Verify Verifiable Credentials submitted to the Organizations (e.g. Service Providers) & (from a user SSI Wallet)
Retail	<ul style="list-style-type: none"> ▶ Issue Service Provider staff with Organization Member credentials ▶ Verify Verifiable Credentials submitted to the Organizations (e.g. Service Providers) & (from a user SSI Wallet)

4.1.3 Threat Landscape

Systems with SSI are prone to experience similar security attacks as any other systems. SSI relies on connectivity and therefore offline use is somewhat limited. Attacks on connectivity and possible denial-of-service attacks may be effective if not mitigated in any way.

Bolgouras et. al [19] have identified the end user (i.e., terminal) device in their architecture to be the most eminent link in terms of security. This is likely to be true in most designs and implementations of SSI. If the end user device is compromised, securing against this kind of security attack requires due diligence from the user or support from the platform (sandboxing applications and fortifying system).

As SSI and for example, granular disclosure, aim at suppressing the disclosed information to bare minimum. One attack against this is the collection of data from multiple sources and combining them by the adversary. Mitigating this requires repudiability of the disclosed information. Mechanisms for such mitigation include undeniable signature which is a digital signature scheme including a disavowal process repudiating the signature. Such disavowal protocol can be included in ZKP. Dominic et. al [20] have published a recent study on ZKP applications. Another option is to use designated verifier signature which is only verifiable for its designated target (one or many). However, these mechanisms do not remove the information possibly gathered earlier on, even though the collector would be unable to verify or prove the origins of the data. In many attack schemes, however, verifying the origin of the data is not required.

The cryptographic components of SSI may be vulnerable to attacks by quantum computers. This can be mitigated by either making sure that quantum-resistant algorithms are used or making the cryptographic algorithms easy to replace in the future.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	60 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

4.1.4 Functional & Non-functional Requirements

Table 17: Self-sovereign identity management Functional Requirements (Author: VTT, UMU, ATOS)

ID	Description	Priority
FR-TRUST-SSIM-001	The TANGO dp-ABC cryptographic system must generate a public and a private key.	H
FR-TRUST-SSIM-002	The TANGO dp-ABC cryptographic system must be able to generate a digital signature.	H
FR-TRUST-SSIM-003	The TANGO dp-ABC cryptographic system must be able to verify signatures with the public key.	H
FR-TRUST-SSIM-004	The TANGO dp-ABC cryptographic system must generate a token corresponding to the selective disclosure ZK proof.	H
FR-TRUST-SSIM-005	The TANGO dp-ABC cryptographic system must be able to verify the Zkey token generated.	H
FR-TRUST-SSIM-006	The TANGO system should give users granular control over the amount of personal information revealed about them, for instance by considering what information is revealed by combining multiple data disclosures made at different times.	M
FR-TRUST-SSIM-007	The TANGO system must provide a wallet or a manager for personal credentials. This is an independent part of the whole SSI module.	H
FR-TRUST-SSIM-008	The TANGO system should provide methods for detecting any use of data contrary to what was agreed to, for instance storing the data for longer time than agreed.	M
FR-TRUST-SSIM-009	The TANGO system could provide methods for repudiation of disclosed data, to minimize the risk of user data being used in harmful ways.	L
FR-TRUST-SSIM-010	The TANGO system must issue eID Verifiable Credentials to end users to simulate the EU Digital Identity Wallet	H
FR-TRUST-SSIM-011	The TANGO system must issue Organization Members with Verifiable Credential to be able to act on behalf of their organization as according to their rôle/ job title.	H
FR-TRUST-SSIM-012	The TANGO system must verify both eID and Organization Member Verifiable Credentials and return the identity claims in an ID Token.	H

Table 18 Self-sovereign identity management Non-Functional Requirements (Author: VTT, UMU, ATOS)

ID	Description	Priority
NFR-SEC-SSIM-001	The TANGO dp-ABC cryptographic system could enable distributed issuance processes, through the aggregation of public keys and combination of signatures into a single one.	L
NFR-SEC-SSIM-002	The TANGO dp-ABC cryptographic system must ensure the confidentiality of sensitive information.	H
NFR-SEC-SSIM-003	The TANGO dp-ABC cryptographic system must ensure the integrity of sensitive information.	H
NFR-SEC-SSIM-004	The TANGO dp-ABC cryptographic system must ensure the authenticity of proofs.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	61 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

NFR-PER-SSIM-001	The TANGO dp-ABC cryptographic system should scale effectively as the number of proof requests and system participants increases.	M
NFR-REL-SSIM-001	The TANGO dp-ABC cryptographic system must produce accurate and correct outputs.	H
NFR-REL-SSIM-002	The system should be highly available, allowing users to access the method without significant downtime or interruptions in case of faults or failures.	M
NFR-COMP-SSIM-001	The SSI Agent must follow the GDPR to respect the privacy of the user's personal data.	H
NFR-SEC-SSIM-005	The SSI Agent must comply with the GDPR and any related regulations as identified by the pilot partners.	H

4.2 Seamless Onboarding for Users and Devices [T4.2]

4.2.1 Description & Functionality

The Onboard SSI is a mobile remote digital onboarding platform for users that are able to provide their physical identity such as passport or eID (electronic Identification), in order to verify their identity and then forward that information to the SSI management component in order to create the relevant Verifiable Credentials. These credentials will then be accessed in order to authenticate to different services. The solution provides a 3-step verification process for the user that performs certain comparisons in order to ensure that the person performing the identity verification process is the same person with the one described on the physical identity. In this way the High Level of Assurance is achieved based on electronic Identification, Authentication and trust Services (eIDAS) regulation. The solution includes artificial intelligence components allowing the system to read through an image the information of the identity document, as well as to crosscheck the facial biometric characteristics of the person performing the onboarding, of the person visible on the identity document and of the person on the NFC chip of the identity document.

- ▶ On-device verification of ICAO electronic Machine Readable Travel Documents such as European member state eIDs and ePassports.
- ▶ On-device biometric capture with face liveness detection verification.
- ▶ On-device biometric verification.
- ▶ Creates an electronic virtual ID, derived from government issued documents including on-device biometric verification against document chip information.

IoT Device SSI component will provide Zero Touch Provisioning (ZTP) of the devices in that when devices are powered on the IoT Device can request directly a public Decentralized Identifier and services can request registration based on their DID automatically. Any subsequent identity verification can be based on DID Authentication. The key functionality supported are as follows:

- ▶ Publish IoT Device public Decentralized Identifier with service endpoint for authentication
- ▶ ZTP of IoT Devices to services
- ▶ Support DID Authentication with challenge /response signed JWTs (See section 4.1)

4.2.1.1 Internal Architecture

The Android app incorporates four modules (Figure 18):

- ▶ Reading Module: Read Machine Readable Travel Documents.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	62 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ Biometrics module.
- ▶ Verification module.
- ▶ Integration module: Interaction with backend on-boarding applications with QR codes.

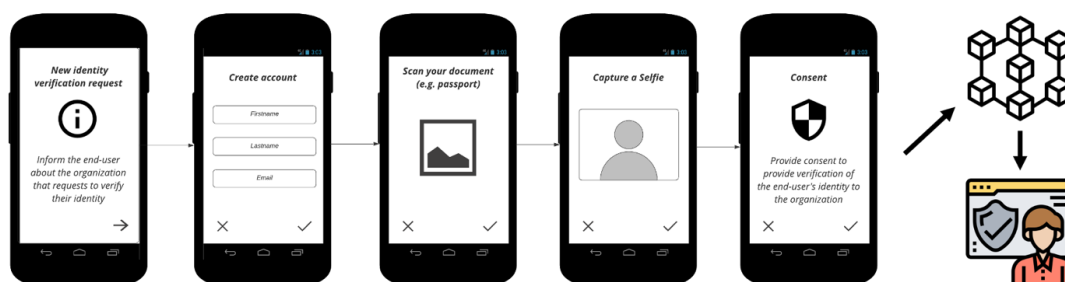


Figure 18: Seamless user onboarding information flow

4.2.1.2 Inputs, Outputs and Interactions with Other Components

The inputs to Onboard SSI component are:

- ▶ End user reading of biometrics and eMRTD (electronic Machine-Readable Travel Document) documents during onboarding.
- ▶ Any user supplied attributes either self-attested or provided by other means e.g., national eID Verifiable Credential.

The output of the Onboard SSI component is the end user's virtual ID.

Onboard SSI for persons and organization members:

The primary interaction of the component is done through the SSI (T4.1) component in order to create verifiable credential by the identity information. This constitutes the initial step of an end-user when they want to join the TANGO platform and create an account. The links created between the TANGO platform and this component are the following:

- ▶ Onboarding request to application with QR Code scan or deeplink either by TANGO Onboarding application (T6.3) or by the Service Provider infrastructure.
- ▶ Result of onboarding with Virtual ID is returned to the link in the QR Code with session information.

IoT Device SSI:

The IoT Device onboarding flow is described below and included in the following figure.

1. Creation of DID upon request by the service logic on the device. The device will create its private key and create a DID Doc which can be hosted on the IoT GW or provider's domain as a DID Web method.
2. The service logic can then proceed to request the device to be registered with a Service Provider with the device's public DID.
3. This triggers a request from the SP to be authorised to the IoT Device SSI for the devices DID.
4. The IoT Device SSI will then send a DID Authentication request to the SP authentication endpoint published in the IoT Devices DID Doc.
5. Once authenticated the IoT Device SSI will return a signed ID Token to the SP.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	63 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

6. The IoT Device is now onboarded with the claims received inside the ID Token.
7. The SP provides an access token to the IoT Device based on the ID Token for establishing secure communications.

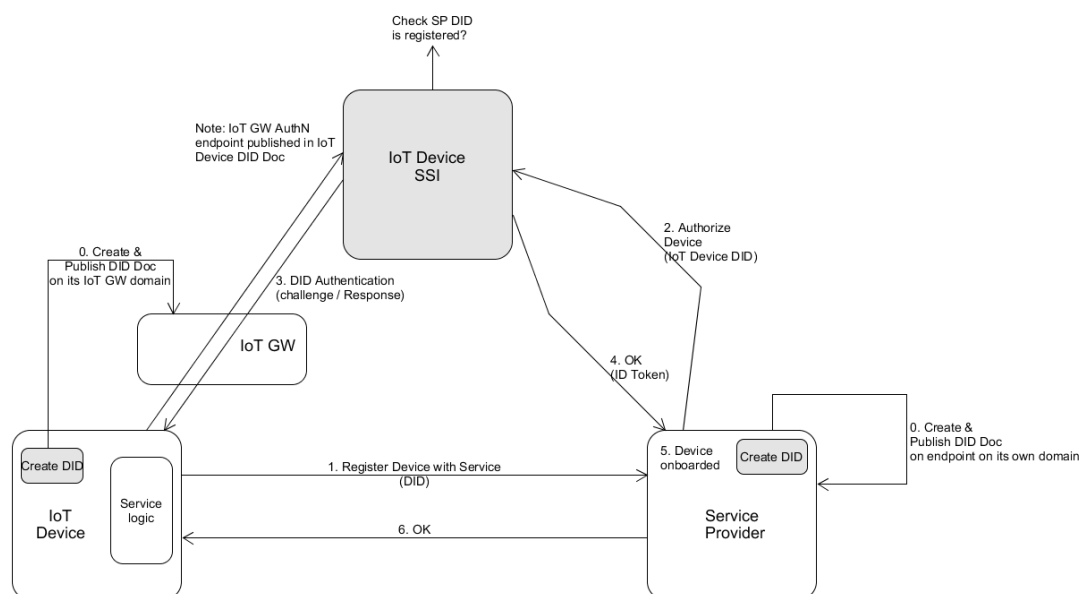


Figure 19: IoT Device SSI component and its interwork for onboarding

4.2.1.3 Significance and Added Value for TANGO

The integration with the TANGO platform will take place through the Task4.1, the Self Sovereign Identity component. The Onboarding component will remotely and securely verify the identity of the user, transforming the identity information into a digital form. The digital form of the identity information will then be sent to the Issuer that will create the Verifiable Credentials.

The Onboarding component will be able to provide secure and easy means for remote onboarding. Currently people are leveraging their existing identity documents such as national id and passport, to perform a transaction (either financial or non-financial). Considering the digital transformation, and the digitalization of data sharing proposed through the TANGO platform, it is essential to securely perform identity verification of users. This component allows end-users to leverage their existing identity documents and transform it to a secure and untampered digital identity, enabling trustworthy transactions.

4.2.2 Associated Pilot Use Cases

Table 19: Association of Pilot Use Cases with SSI Onboard component of T4.2

Pilot	Association
Smart Hospitality	Guests will be able to perform remote identity verification before they arrive at the hotel, thus they may not have to provide their identity information to the reception.
Autonomous Vehicles	Drivers and other relevant users will be able to verify their identity in order to generate the relevant verifiable credentials, providing access to the vehicle and other services in the vehicle such as personalised infotainment. The component will provide the ability for onboarding of devices.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	64 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Smart Manufacturing (FMAKE & RIAS)	<p>Case 1: Staff members will verify their identity seamlessly and then ensure that the customers access only the necessary information.</p> <p>Case 2: Staff members will verify their identity seamlessly in order to allow them to manage their data and access the necessary information in terms of the factory and the customers.</p> <p>The component will provide the ability for onboarding of devices.</p>
Public Administration	Applicants will be able to perform identity verification remotely in order to begin the process and apply for a visa remotely.
Retailers	Customers and employees from the retailer will perform identity verification, creating verifiable credentials, in order to use them for accessing their data including reports and other relevant information.

4.2.3 Threat Landscape

The known threats and attacks that may occur to the particular component are related to the falsification of the identity. These attacks are split into the following:

- ▶ Fake identity document. The attacker in this case will provide a fake identity document, where the eMRTDs are fake or stolen.
- ▶ Presentation attack. The attackers in this case may use a fake/legit identity document in order to verify their identity and will attack the system by presenting another person's facial biometric through face spoofing (photograph, video) and/or deep fakes.

The component incorporates specific threat detection mechanisms in order to be able to prevent cases of fake identity documents as well as presentation attacks where the attacker tries to mimic a person's facial image. The following defence mechanisms are employed in the component:

- ▶ Fake identity document. The component verifies the identity information of the identity document by comparing the information extracted through the MRZ (Machine Readable Zone), the readable information that is visible on the document which is scanned through Optical Character Recognition and the information retrieved through the NFC-chip (Near-Field Communication), providing High Level of Assurance. According to the eIDAS regulation, these constitutes the different steps in order to verify the identity of a person based on a physical identity document. This way allows a three-step check of the validity of the document.
- ▶ Presentation attacks. The component incorporates a liveness detection mechanism that allows the detection of various presentation attacks such as paper image, image on a screen, a video, 2-D mask and potentially 3D-masks.

4.2.4 Functional & Non-functional Requirements

Table 20: Seamless Onboarding for Users and Devices Functional Requirements (Author: QBE)

ID	Description	Priority
FR-TRUST-SOUD-001	The TANGO system must allow hotel guests to perform remote identity verification.	H
FR-TRUST-SOUD-002	The TANGO system must allow visa applicants to perform remote identity verification.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	65 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Table 21: Seamless Onboarding for Users and Devices Non-Functional Requirements
(Author: QBE)

ID	Description	Priority
NFR-COMP-SOUD-001	The TANGO system must be compatible with SSI.	H
NFR-PER-SOUD-001	The TANGO system must support identity verification in a performant, reliable, secure manner.	H
NFR-EXIST-SOUD-001	The TANGO system must perform the identity verification through smartphones.	H

4.3 User continuous behavioural authentication [T4.3]

4.3.1 Description & Functionality

User Continuous Behavioural Authentication – UCBA Asset

UCBA is an Android mobile technology allowing the continuous assessment of the user, when they are accessing a particular mobile app, through learning their behavioural patterns related on how the user is interacting with the device for example on how the user holds the device, the way they perform swipes, gestures and typing behaviour. UCBA constitutes a security mechanism that provides high reassurance to the user and the administrators of a system, that only the account owner has access to the system. Considering that no action is required from the user as the system processes information in the background, UCBA offers better user experience as they do not have to perform any action to authenticate themselves, in comparison to existing cumbersome 2-factor authentication that creates huge frustration to the users. The following features are provided from UCBA technology:

- ▶ Derives behavioural characteristics of the user related to how the user is moving, how the user is interacting and using the device and transactional patterns.
- ▶ Offers on-device and cloud-based biometric authentication for fingerprint and facial recognition.
- ▶ Combines the aforementioned behavioural traits and biometrics through a novel AI-based fusion mechanism.
- ▶ Optional, could provide input into the attribute-based authentication at the sticky policies.
- ▶ Integrates with existing mobile apps such as wallets through an SDK.

4.3.1.1 Internal Architecture

The Android app incorporates four modules:

- ▶ Security and storage module for managing the persistent local variable and information.
- ▶ Data collection module that is responsible for collecting all the sensor, device and transactional data required for the processing and inference.
- ▶ Inference module that performs certain pre-processing and inference on-device before offloading the data to the cloud
- ▶ Communication module that sends the information to the cloud platform and is responsible for the secure communication of the two components.
- ▶ Integration module: Interaction with mobile applications through the SDK interface.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	66 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

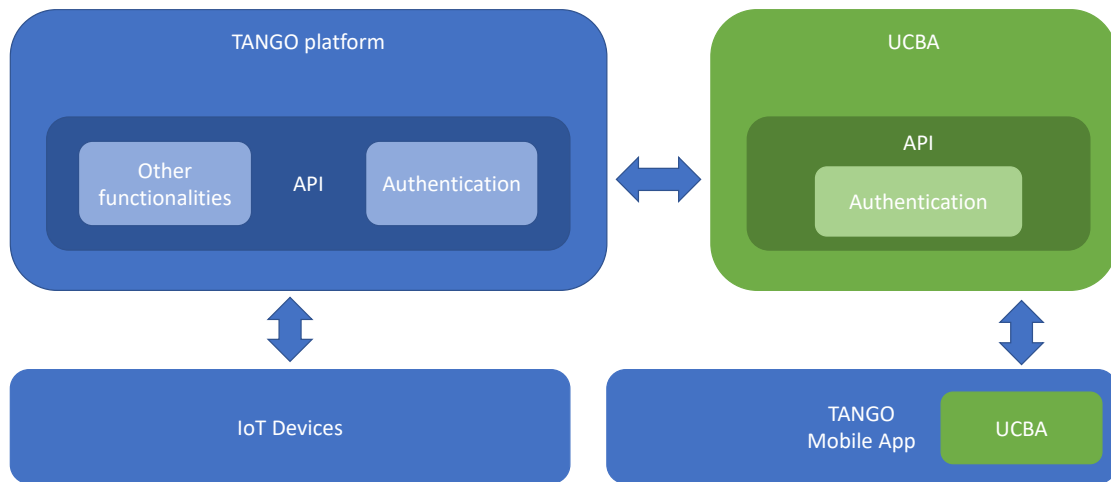


Figure 20: Information flow for the different components involved regarding the UCBA component

4.3.1.2 Inputs, Outputs and Interactions with Other Components

The inputs to the UCBA component are: Sensor data from the device such as inertial data (accelerometer, gyroscope, magnetic field), geolocation data (GPS, WiFi, Bluetooth), screen data (touch events, gestures), camera (image frames), fingerprint (on-device biometric), activity (walking, running, riding a bicycle etc.), screen activity (at which screen of the app is the user), transaction (beneficiary, amount, time, location), device information. The outputs of the UCBA component are: The user status (legit, imposter, unavailable) and the confidence level. (low, medium, high, very high).

The primary interaction of the component is done through the TANGO Wallet introduced in T6.3 and potentially T4.1 component in order to provide high security to access for the particular wallet app. The T4.3 performs secure authentication when the user logs in and throughout the whole session while accessing the TANGO Wallet, in order to ensure that only the legitimate user has access. Furthermore, collaboration with the T3.3 - Confidentiality and Privacy by Design is also explored to feed the behavioural risk score into the attribute-based authentication i.e., the sticky policies.

4.3.1.3 Significance and added value for TANGO

The integration with the TANGO platform will take place through the following two components:

- ▶ The mobile wallet app that will be introduced in the TANGO platform which will be responsible for managing the verifiable credentials of the end-user. This mobile app requires high level of security to manage the credentials and verify the various parts of the users' identity.
- ▶ The core authentication mechanism of the TANGO platform. This mechanism has its roots on attribute-based authentication and sticky policies, where the behavioural risk score will provide additional input to these policies and incorporate the behavioural element as an extra authentication mechanism.

4.3.2 Associated Pilot Use Cases

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	67 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 22: Association of Pilot Use Cases with UCBA component of T4.3

Pilot	Association
Smart Hospitality	Guests will be able to access the mobile app of the hotel, in order to share information with the hotel, access services and perform transactions.
Smart Manufacturing (FMAKE)	Workers will be continuously authenticated when using the printer jobs, when accessing a particular infrastructure and for remote consultants that conduct fixes to the infrastructure.
Public Administration	Strong authentication for applicants, employees, employers when accessing the TANGO platform.
Banking	Continuous behavioural authentication incorporates transaction patterns that allow the detection of unauthorised access and potential fraud. The solution can provide the basis for the transactional patterns detection in a federate learning concept. Further behavioural elements such as the device usage, biometrics and user movements can be potentially explored.

4.3.3 Threat Landscape

The known threats and attacks that may occur to the following attacks:

- ▶ Impersonation attacks: These attacks occur by a potential attacker that is trying to mimic the behavioural patterns of a legitimate user that is striving to get access to the system. An example of this could be done by logging the user's behavioural data and then leveraging ML mechanisms to create synthetic behavioural data.
- ▶ Replay attacks: In this attack, the attacker is logging the legitimate user's behavioural patterns and then replays them later on in order to gain access to the system. Usually, this process is done by intercepting and recording the network traffic.
- ▶ Evasion attacks: The attackers are deliberately modifying their behaviour in order to bypass the authentication system, using techniques such as randomisation, obfuscation and noise injection, to avoid detection of anomalies.
- ▶ Model poisoning attacks: Attackers are injecting malicious data such as false data or modification of existing data into the system to confuse the ML models.

The component incorporates specific threat detection mechanisms that allow the filtering and validation of the data at various stages of the inference process. It is essential to filter and cross-check the data that are provided by the smartphone device to understand if the device is malfunctioning or has been tampered. To this end, the device behavioural authentication mechanism is also important to ensure that it has not been tampered. Considering that the UCBA is leveraging data directly from the operating system of the device, provides more security than requiring the user to provide input. However, as mentioned it is important to ensure that the device of the user is not tampered or malfunctioning.

4.3.4 Functional & Non-functional Requirements

Table 23: User continuous behavioural authentication Functional Requirements (Author: QBE)

ID	Description	Priority
FR-TRUST-UCBA-001	The TANGO system must guarantee that the hotel guest accessing the platform, is who they claim they are.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	68 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

FR-TRUST-UCBA-002	The TANGO system must guarantee that the person (staff employee) accessing the platform, is who they claim they are.	H
FR-TRUST-UCBA-003	The TANGO system must guarantee that the person (visa application and organisation employee) accessing the platform, is who they claim they are.	H
FR-TRUST-UCBA-004	The TANGO system must guarantee that the person accessing the platform, is who they claim they are.	H
FR-TRUST-UCBA-005	The TANGO system must be able to detect fraud in different transactions.	H

Table 24 User continuous behavioural authentication Non-Functional Requirements (Author: QBE)

ID	Description	Priority
NFR-COMP-UCBA-001	The solution must be compatible with the TANGO mobile and web apps.	H
NFR-REL-UCBA-001	The TANGO system must ensure that the user continuous behavioural authentication will operate in real-time, reliably and to be secure.	H
NFR-EXIST-UCBA-001	The TANGO system must support smartphone and web apps	H
NFR-REL-UCBA-002	The TANGO system should be able to accurately and reliably detect fraud based on different types of transactions.	M

4.4 Device continuous behavioural authentication [T4.4]

4.4.1 Description & Functionality

This component will provide all the required functionality for the continuous authentication of a user's device via the online behavioural analysis and learning of the unique patterns and characteristics of a device's operation such as its hardware configuration, software settings, usage patterns, and network activity. Through the analysis of the aforementioned patterns and characteristics, a baseline is determined, and a continuous monitoring mechanism compares it to the established baseline. Abnormal device behaviour detection, classified by the trained AI and ML models, such as change in device's location or an unusual network connection, will trigger alerting mechanisms that will notify the subscribers of this service/component for authentication failure events. The features to be covered from this component are the following:

- ▶ Online behavioural analysis and learning to identify unique patterns and characteristics of a device's operation.
- ▶ Analysis of hardware configuration, software settings, usage patterns, and network activity to establish a baseline for normal device behaviour.
- ▶ Continuous monitoring of device behaviour and comparison to the established baseline.
- ▶ AI and ML models that detect abnormal device behaviour,
- ▶ such as changes in location or unusual network connections and traffic patterns.
- ▶ Alerting mechanisms that notify subscribed services of authentication failure events.

4.4.1.1 Internal Architecture

The Device Continuous Behavioural Authentication component consists of the following pipeline with its sub-components interacting in a sequential order:

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	69 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ **Data extraction agent:** This sub-component is responsible for monitoring, polling, logging and streaming the device's hardware info (id, manufacturer and model, processor type and speed etc.), hardware configuration (screen resolution, etc.), network status (IMEI number, IP address, MAC address, WLAN SSID, etc.), network traffic patterns (DNS queries, Packet size and timing, port usage, etc.) and location familiarity (nearby RF signals e.g. WiFi, Bluetooth devices) using state-of-the-art data streaming software and protocols such as Apache Kafka and MQTT.
- ▶ **Data pre-processing module:** This sub-component is responsible for rendering the extracted data in a pre-defined and standard format, also taking care for handling invalid entries and/or undefined values. The processed data is then fed to ML models for behavioural analysis and classification.
- ▶ **AI/ML model for behavioural analysis:** This sub-component will include a fitted AI model for behaviour recognition and anomaly detection. Several algorithms will be considered and evaluated in terms of traditional classification metrics such as accuracy, precision, f1 score and Area Under the Curve (AUC). The algorithms to be considered are Random Forest, Hidden Markov Models (HMMs), Recurrent Neural Networks (RNNs) including Long Short-Term Memory (LSTM) Networks and Convolutional Neural Networks (CNNs). This sub-component will rely on the interaction with the AIaaS and MLOps components of T5.2.
- ▶ **Alert notification module:** This sub-component takes as input the verdict of the AI model for abnormal behaviour and generates alerts that provide contextual information such as the affected device and the details of behaviour change that triggered the alert. For the latter, explainable AI techniques will be leveraged in order to interpret AI model's output. The alerts will be forwarded to further trust management TANGO components with the use of message protocols and queues such as XMPP and MQTT.

The *data extraction agent* will be deployed on the users' devices, while the *data preprocessing* module, the *AI/ML models* and the *alert notification module* will be deployed as microservices on the cloud/edge.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	70 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

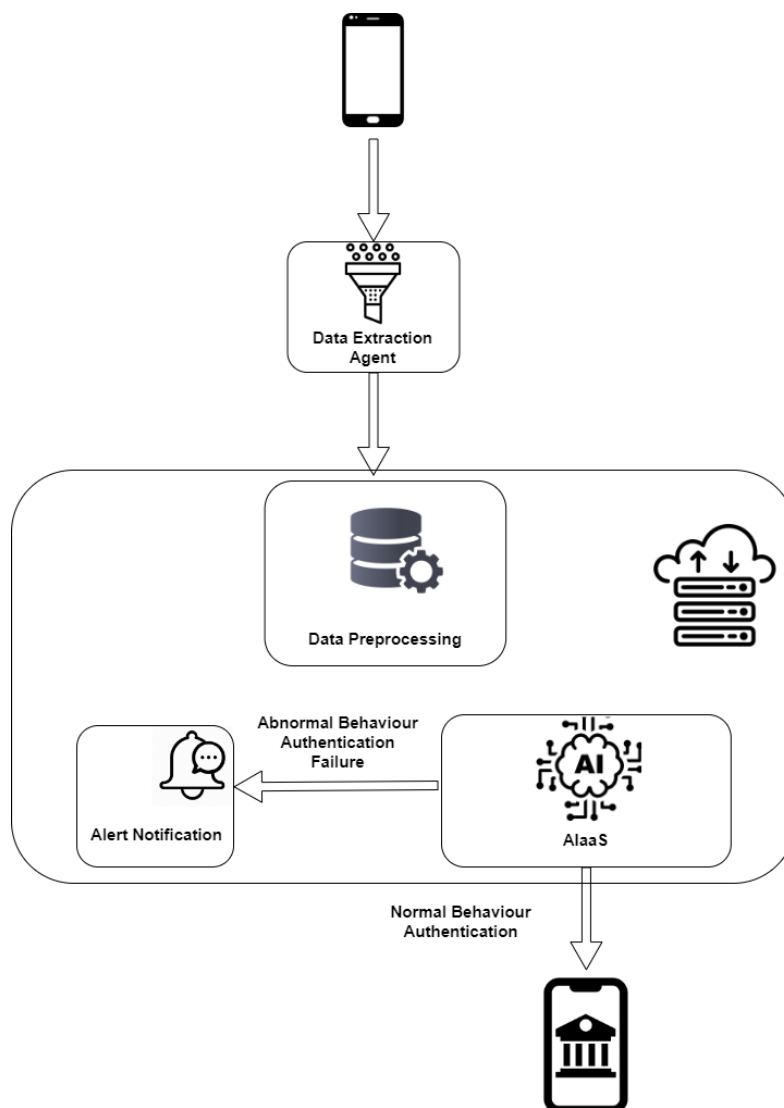


Figure 21: Device Continuous Behavioural Authentication component's architecture

4.4.1.2 Inputs, Outputs and Interactions with Other Components

This component takes as input the device's hardware info (id, manufacturer and model, processor type and speed etc.), hardware configuration (screen resolution, etc.), network status (IMEI number, IP address, MAC address, WLAN SSID, etc.), network traffic patterns (DNS queries, Packet size and timing, port usage, etc.) and location familiarity (nearby RF signals e.g. WiFi, Bluetooth devices).

This component provides as output alerts that indicate abnormal behavior detection along with contextual information about the violating pattern.

The Device Continuous Behavioural Authentication component interacts with:

- ▶ The Trustworthy Data Sharing (T3.2) component for the distributed tokenization of data for the exchange of monitored data between the devices and the cloud/edge services.
- ▶ The Confidentiality and Privacy component (T3.3) for setting data expiration time.
- ▶ The Self-Encryption and Decryption component (T3.4) for the encrypted transmission of monitored data.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	71 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ The Seamless Onboarding for Devices (T4.2) for creating the identifiers based on device characteristics.
- ▶ The Exploratory Data Analysis Engine (T5.1) for the data handling and transformation of the *data pre-processing module*.
- ▶ The Energy-Efficient AI model training (T5.2) for feature selection and training of AI models.
- ▶ The X-AI tools (T5.5) for providing interpretations of AI model classification of abnormal behaviour.
- ▶ The administrative web interfaces (T6.3) and visualization tools (T6.4) for alert notification.

4.4.1.3 Significance and Added Value for TANGO

The Device Continuous Behavioural Authentication component will provide the means for the TANGO platform to ensure the trustworthiness of data exchanged between the participating stakeholders by providing an automated, continuous behavioural authentication mechanism that prevents the compromise and exploitation of IoT devices.

4.4.2 Associated Pilot Use Cases

Table 25: Association of Pilot Use Cases with DCBA component of T4.4

Pilot	Association
Autonomous Vehicles	In the Autonomous Vehicles use case, the Device Continuous Behavioural Authentication (DCBA) module will provide the means of ensuring that the device of the user (smartphone, tablet) is always authenticated before issuing any request (booking of a “robo-taxi”, an early stopping of the car while on the move, etc.) or receiving data from the car and/or the server (car location, speed, ETA, etc.). Specifically, after the device has been authenticated through the SSI procedure, the DCBA component will be constantly monitoring and polling several metrics and info such as hardware info, network status, network traffic patterns, power consumption, etc. and comparing them through the application of a Machine Learning model to a baseline “normal” behaviour in order to detect if there is a deviation from it and thus a new SSI authentication procedure is needed.
Smart Manufacturing – Case 2 (RIAS)	Regarding the second use case of Smart Manufacturing, the DCBA module will be employed in order to constantly monitor the wearable devices of the shopfloor workers (smartphones, smartwatches, tablets, etc.) and ensure that they are moving into the designated areas for which they are authorized. Specifically, the DCBA component will monitor, besides GPS location that may be unavailable at times, several network info and metrics such as neighbouring WLAN SSIDs, received signal strength (RSSI), and network performance metrics (packet loss percentage, airtime utilization, etc.) in order to be able to localize the device within the shop floor area. In the case of entering an area for which the worker has no authorization, the DCBA alert notification mechanism will generate an informative alert that can be visualized in the respective smartphone application of the shopfloor supervisor.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	72 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

4.4.3 Threat Landscape

The known threats of continuous behavioural authentication mechanisms that are relevant to the Device Continuous Behavioural Authentication component are:

- ▶ Impersonation attacks: Impersonation attacks involve an attacker mimicking the behaviour patterns of an authenticated device to bypass the continuous authentication mechanism.
- ▶ Data poisoning attacks: Data poisoning attacks involve an attacker intentionally modifying or injecting malicious data into the behavioural data used for authentication, which can lead to incorrect or false positive authentication results.
- ▶ Man-in-the-middle attacks: Man-in-the-middle attacks involve an attacker intercepting and altering the communication between a device and the continuous authentication system, which can lead to incorrect or compromised authentication results.
- ▶ Breach of privacy: The collection and analysis of user behaviour patterns for authentication purposes can lead to potential breaches of privacy if the data is not properly secured or anonymized.

For mitigation of the aforementioned security risks the Device Continuous Behavioural Authentication component will rely on other components being developed during the TANGO project. Specifically:

- ▶ Multi-factor authentication: the combination with the User Continuous Behavioural Authentication component (T4.3) will assist in strengthening the security of the authentication system and reduce the risk of impersonation attacks.
- ▶ Data encryption and trustworthy data sharing (T3.4 and T3.2 respectively) will help to prevent data poisoning attacks by ensuring that data is only accessible to authorized personnel and protected from unauthorized access or modification.
- ▶ Blockchain-based and trustworthy data sharing (T3.1 and T3.2 respectively) will assist in preventing man-in-the-middle attacks.
- ▶ Confidentiality and Privacy component (T3.3) and trustworthy data sharing (T3.2) will minimize the risk of privacy breaches by allowing owners of data to determine who has the right to access it as well as setting expiration time, while distributed tokenization of data will preserve the privacy of exchanged sensitive/personal data.

4.4.4 Functional & Non-functional Requirements

Table 26: Device continuous behavioural authentication Functional Requirements (Author: UTH)

ID	Description	Priority
FR-TRUST-DCBA-001	The TANGO system must be able to monitor, poll and log the device's hardware information (id, manufacturer, model, processor type, speed, etc.)	H
FR-TRUST-DCBA-002	The TANGO system must be able to monitor, poll and log the device's network status (IMEI number, IP address, MAC address, WLAN SSID, etc.) as well as the device's network traffic patterns (DNS queries, packet size and timing, port usage etc.)	H
FR-TRUST-DCBA-003	The TANGO system must be able to monitor, poll and log the device's location familiarity (nearby RF signals such as WiFi, Bluetooth devices, etc.)	H
FR-TRUST-DCBA-004	The TANGO system should be able to render extracted data in a predefined and standard format, handle invalid entries and undefined values and prepare the processes data for	M

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	73 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

	feeding into AI/ML models for behavioural analysis and classification.	
FR-TRUST-DCBA-005	The TANGO system must be able to feed the preprocessed data to an AI/ML model for behavioral recognition and anomaly detection.	H
FR-TRUST-DCBA-006	The TANGO system must be able to generate alerts based on the verdict of the AI model for abnormal behavior, and provide contextual information, including the affected device along with details of the behavior change that triggered the alert.	H

Table 27: Device continuous behavioural authentication Non-Functional Requirements
(Author: UTH)

ID	Description	Priority
NFR-COMP-DCBA-001	The TANGO system should ensure compatibility and applicability of the DBCA module.	M
NFR-PER-DCBA-001	The TANGO system should ensure the efficient execution of the DBCA module, as well as the rapid reaction to abnormal behavior.	M
NFR-REL-DCBA-001	The TANGO system must ensure continuous and accurate device behavioral authentication.	H
NFR-SEC-DCBA-001	The TANGO system must ensure that personal data is not leaked, and data handling is GDPR-compliant.	H

4.5 Hardening Against Side-channel Attacks [T4.5]

4.5.1 Description & Functionality

The provided component will be a compiler able to automatically apply side-channel attack countermeasures. We will consider code polymorphism and loop shuffling countermeasures. Code polymorphism consists of making the sensitive code vary dynamically using diverse transformations such as the random insertion of noise (useless) instructions, random permutation of the registers, the random replacement of an instruction by a semantically equivalent sequence of instructions, or the random shuffling of independent instructions. Loop shuffling consist in executing loop iterations in a random order when possible.

In TANGO, code polymorphism countermeasure will be extended with new transformations to enhance the security brought by the countermeasure and broaden the configuration space. Support for random loop shuffling will also be added to our compiler, a countermeasure that has a low performance overhead. The compiler will be able to apply both countermeasures combined. It will support the ARMv7m architecture, and it will be used to produce side-channel resistant code.

4.5.1.1 Internal Architecture

Our compiler is based on LLVM (<https://llvm.org/>), with clang front-end (<https://clang.llvm.org/>). The compiler is structured as follows: clang parses the source code and generates an intermediate representation called *LLVM IR*, then the middle-end performs code analyses and code transformations on LLVM IR, then the back-end transforms the LLVM IR into assembly or object code, with some target-specific code transformations and optimisations. Each code analysis or code transformation is called a *pass*. The compilation flow is thus composed of a suite of passes.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	74 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

LLVM is modified in order to add passes that apply the countermeasures. While code polymorphism is applied during code emission in the back-end, loop shuffling will probably be implemented in the middle-end, on LLVM IR.

Note that the compiler itself will not be integrated in a system but will be used to protect the functions necessary for the system: the source code files that contain sensitive functions will be built using the modified compiler, and the obtained object files with countermeasures will be integrated in the system.

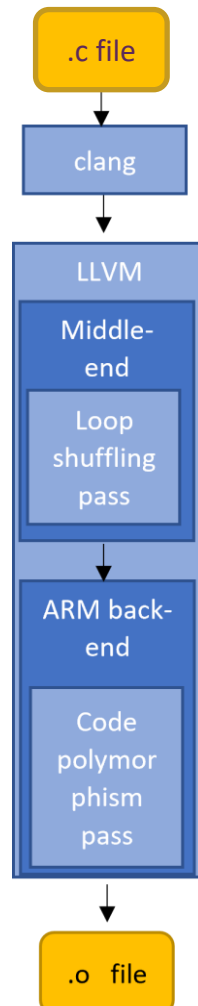


Figure 22: Compiler's internal architecture

4.5.1.2 Inputs, Outputs and Interactions with Other Components

The inputs to the compiler are the C file containing the functions to be secured against side-channel attack; Annotations that specify the functions to be secured, and potentially other information (like the loops to shuffle); In addition, configuration of the countermeasures, e.g., which transformations to enable for code polymorphism.

The output of the compiler is an object or assembly file, for ARMv7m, with countermeasures applied on the specified functions.

The compiler will be used on the cryptography functions (e.g., AES) used e.g., for communication purposes. It will be used in order to add countermeasures against side-channel attacks for these functions. The compiler will take a source code containing the considered cryptography functions and generate an object file with the countermeasures enabled.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	75 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

4.5.1.3 Significance and Added Value for TANGO

The technology represents a key feature for the project wherever side-channel attack can be a threat. Indeed, side-channel attacks have a high ability to find cryptographic keys used in embedded systems or IoT device. Such keys are a critical component as they can be used to secure data communication, as well as firmware updates. Hence, such systems need to be hardened against side-channel attacks.

4.5.2 Associated Pilot Use Cases

Table 28: Association of Pilot Use Cases with LLVM-based compiler of T4.5

Pilot	Association
Smart hospitality	In this use-case, smart sensors in the user's room communicate measurements to a server. An attacker could have access to the sensors and the associated embedded board and could find an encryption key through a side-channel attack. The encryption key can be used for various purposes: denial of service, data stealing, data tampering, authenticating a malicious firmware update etc. The compiler will be used to harden the crypto implementation used against side-channel attacks to prevent this threat.

4.5.3 Threat Landscape

Advanced machine learning attacks with strong attacker expertise have been shown to threaten the code polymorphism countermeasure [21]. It is planned to combine the code polymorphism with loop shuffling to make attacks harder, as well as to deploy new code transformations inside code polymorphism in order to strengthen the countermeasure. Finally, the high configurability of the countermeasure allows to choose stronger configuration than the one analysed in the aforementioned publication.

4.5.4 Functional & Non-functional Requirements

Table 29: Hardening Against Side-channel Attacks Functional Requirements (Author: CEA)

ID	Description	Priority
FR-TRUST-HASA-001	The smart hospitality components for sensor data ciphering physically accessible to an attacker must include side-channel countermeasures.	H

Table 30: Hardening Against Side-channel Attacks Non-Functional Requirements (Author: CEA)

ID	Description	Priority
NFR-EXIST-HASA-001	The smart hospitality components for sensor data ciphering physically accessible to an attacker must include side-channel countermeasures.	H
NFR-COMP-HASA-001	The code polymorphism countermeasure relies on dynamic code generation: write and execute permissions should be given to a memory segment in order to deploy this countermeasure.	M
NFR-SEC-HASA-001	Securing against side-channel attacks forces to find a trade-off between performance (i.e. clock cycles) and security (i.e.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	76 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

	how resistant to an attack the code will be). Security must be as good as possible for a given performance constraint.	
--	---	--

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	77 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

5 AI-based Framework for Green & Trustworthy Operations

5.1 Exploratory Data Analysis Engine [T5.1]

5.1.1 Description & Functionality

The exploratory data analysis engine (EDAE) will be fed with data deriving from the end users aiming to perform data analytics with a focus on the identification of patterns or features of interest among the data. EDAE will be capable of: a) identifying correlations among the heterogeneous real-world data that will be inserted into the system and b) selecting correlation-based features. A combination of well-established data analysis techniques will be exploited in order to develop an AI-driven method to automatically perform the correlation analysis, visualising the generated results. The module's envisioned process consists of the following steps: 1) Variable Identification, 2) Uni-variate Analysis, 3) Bi-variate Analysis, 4) Missing values treatment, 5) Outlier treatment, 6) Variable transformation and 7) Variable creation and selection.

EDAE will implement some of the following functionalities to prepare/analyse the data received:

- ▶ Univariate Analysis or Bi-variate Analysis
- ▶ Deletion, Mean/Mode/Std/Median value, Prediction model, KNN Imputation
- ▶ Rows deletion, binning values transformation, imputation, separate treatment
- ▶ Logarithm, Square/Cube root, Binning transformation
- ▶ One hot encoding
- ▶ Principal Component Analysis (PCA), Uniform Manifold Approximation and Projection (UMAP), Linear regression, Forward/Backward/Stepwise selection, Random Forest, Lasso Regression
- ▶ Clustering techniques, Matrix Profile

5.1.1.1 Internal Architecture

The EDAE consists of 3 main sub-modules: 1) A data analytics sub-module that is responsible for studying the relationship between variables among the data provided by generating descriptive statistics, data type summary, extracting the data characteristics, etc.; the sub-module will be also capable of performing univariate, bivariate and multivariate analysis. 2) The second sub-module is responsible of handling missing or duplicate data, providing the option to the user to decide how the platform will treat this data or asking him to confirm whether they would like to proceed with the proposed treatment methods. 3) The third sub-module aims to prepare the data to be used by other TANGO components in order to train their AI/ML algorithms. The functionalities of this sub-module include the semi-automated data transformation, the implementation of dimensionality reduction techniques (AI may also be implemented to assess what is the optimal dimensionality), variable selection and feature engineering as well as outlier identification and treatment, if needed and upon confirmation by the user.

The methods/techniques that have been implemented as part of the EDAE tool have been developed in Python 3, and all the references to different libraries or functions are correlated with the Python programming language. In the next paragraphs, more details about the implementation of each sub-module are presented.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	78 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

The results of the **data analytics sub-module (1st sub-module)** inform the reasonable next steps you can take with the data, what corrective actions you will need to perform, and generally help better understand the range of values and data types that you will be working with during a machine learning project.

Under descriptive analysis, there have been created a number of functions to display useful characteristics of the data. These characteristics include:

- ▶ **The 'basic characteristics' of the DataFrame:** the number of rows and columns in the raw data, and how much space it takes in memory.
- ▶ **Missingness analysis:** the number and proportion of missing values in each column in the data set. (Missingness is further explored in the Missing Value Handling section)
- ▶ **Data type summary:** the inferred data type of each column. Here, we aggressively try to identify each data type, including date-time variables (although this can be turned off while using the function)
- ▶ **Descriptive statistics:** a table of the mean, minimum, maximum values in each column as well as the 25th, 50th and 75th percentiles of the data (clearly, this only applies for numeric variables)
- ▶ **Duplicate and constant analysis:** the number of rows and/or columns that are duplicates of another row or column, and all columns that contain a constant value.
- ▶ **The descriptive analysis functions** also include the ability to plot the numeric and categorical variables in the data. Histograms are plotted for the numeric variables, to identify their overall distribution; frequency charts are created for categorical variables (when less than ten distinct categories are present).

Additionally, a separate set of functions further explores the behaviour of variables in the data set and their relationships to one another. Under the univariate analysis functions, plots of the kernel density estimations for each of the numeric variables are generated to better understand the underlying distribution function that may be generating the observed variables. Under the multivariate analysis functions, the correlation matrix between the variables is printed (using Pearson correlation). Plots with the relationship between all continuous variables and each date time variable in the data set are also generated. The idea here is to help the user visually identify any patterns in the numeric variables that emerge over time.

The **2nd sub-module** is responsible for **handling missing data**. Missing values are almost inevitable in any reasonably sized data collection effort. Many algorithms fail in the presence of missing values, and they must be dealt with in some way or another. Values can be missing from the data set for a variety of reasons: it can be random, or it can correlate other values in the data or characteristics of the data subject. EDAE do not deal with the different reasons for why data can be missing in this set of functions. Instead, EDAE presents a further set of charts to better explore the missingness in the data set, as well as a variety of methods to handle them. Three charts are plotted to visualise the missing values in the data set:

- ▶ The *missingno matrix*: presents a bird's eye view of the missing data in the data set, and allows for comparisons to be made between columns and identify the rows with the highest levels of missingness
- ▶ The *missingno bar chart*: shows how much data is present as a percentage in each column
- ▶ The *missingno heat map*: shows the missingness correlation between columns; this is, it shows which columns are more likely to both have missing values, which have a negative relationship and which columns' missing values are independent of one another.

The techniques provided to handle missing values are as follows:

- ▶ **delete:** remove all rows that have a missing value.
- ▶ **mean:** fill all missing values with the mean of that column.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	79 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

- ▶ median: fill all missing values with the median of that column.
- ▶ mode: fill all missing values with the mode of that column.
- ▶ K nearest neighbours: identify the nearest 1% of the data set and fill missing values with the mean of the nearest neighbours

The **3rd sub-module** is responsible for several tasks that are presented below. Overall, it aims to prepare the data to be used by other TANGO components in order to train their AI/ML algorithms.

Outlier Treatment:

Outliers are observations/values that appear far away from other observations and diverge from the overall pattern in a sample. Outliers can be classified into two categories:

- ▶ Univariate: based on the distribution of a single variable
- ▶ Multivariate: based on the multivariate distribution of two or more variables

Some examples of outliers are:

- ▶ Any data point/value which is beyond the range of $\pm 1.5 * \text{IQR}$ (interquartile range).
- ▶ Any data point/value out of the range 5th ÷ 95th percentile.
- ▶ Any data point/value that is three or more standard deviations away from mean.

When dealing with outliers two steps are considered, namely *outlier identification* and *outlier treatment*.

For *outlier identification*, the most commonly used methods to detect outliers are visualization methods as well as other statistical methods and algorithms. Boxplots, Histograms, and Scatter Plots can be successfully used as visualization methods. Z-score, interquartile range (IQR) are two statistical methods that can be used, and Isolation Forest is an algorithm that can help identify outliers.

For *outlier treatment*, a range of methods/techniques that can be used. For example, the detected outliers can be removed if it is due to data entry error, can be imputed with median, mean, standard deviation imputation methods or a machine learning model (e.g., Linear Regression, Random Forest) can be used to impute the detected outliers with some predicted values. Pre-processing techniques can also be applied, such as PowerTransformer() or RobustScaler() (all from scikit-learn) to reduce the impact of the detected outliers. EDAAE is intended to detect and treat outliers using the methods mentioned above.

Dimensionality Reduction:

Dimensionality reduction is the transformation of data from a high dimensional space into a low dimensional space so that the low dimensional representation retains some meaningful properties of the original data. Dimensionality reduction can be used to reduce the noise in the data, for data visualization by eliminating the redundant features, to help the improvement of the model's accuracy/performance, to reduce the computation time by 'compressing' the data which leads to faster training of the machine learning models.

There are multiple methods for reducing the number of dimensions in a data set, and some of them have been integrated into the EDA tool. These methods include *Principal Component Analysis (PCA)*, *Uniform Manifold Approximation and Projection (UMAP)*, and *t-Distributed Stochastic Neighbour Embedding (t-SNE)*. Each of these techniques employs a unique approach to map the data onto a lower-dimensional space while preserving important information.

Variable Importance/Feature Selection:

Variable importance can only be determined in the context of a prediction problem. It is a means of determining which columns in the data set contribute the most to the performance of the model. When the variable importance function is called from the EDA tool, three methods are

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	80 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

employed to calculate variable importance. The variables the methods had in common are printed on a final line. To use this functionality, a target variable must be specified against which the other variables are ranked.

The three methods of determining variable importance are:

- ▶ **Random Forest:** Random forests can determine the importance of a variable by measuring the average decrease in the loss function each variable contributes. The variable with the largest average decrease is considered the most important, and the remainder are ranked accordingly.
- ▶ **Variable Permutation:** In this method, each column, one at a time, is randomly shuffled so that the observed values no longer belong to the original observation. A model is fit between this new data and the target variable and compared to the original baseline model (in our case, the random forest above). If the variable is indeed important, then shuffling the value should lead to a decrease in the performance of the model. If there is no relationship between the variable and the target, then shuffling the observation should have no effect on model performance. The shuffled variables that lead to the largest drop in performance are those considered most important in the model.
- ▶ **SHapley Additive exPlanations (SHAP):** SHAP is an approach that tries to estimate the contribution of each variable to the prediction outcome. The magnitude of the Shapley value is proportional to the contribution (positive or negative) that the variable makes to the outcome. The average of the absolute value of the Shapley values for each column gives a way to rank each variable in terms of its contribution and thus importance in the prediction problem.

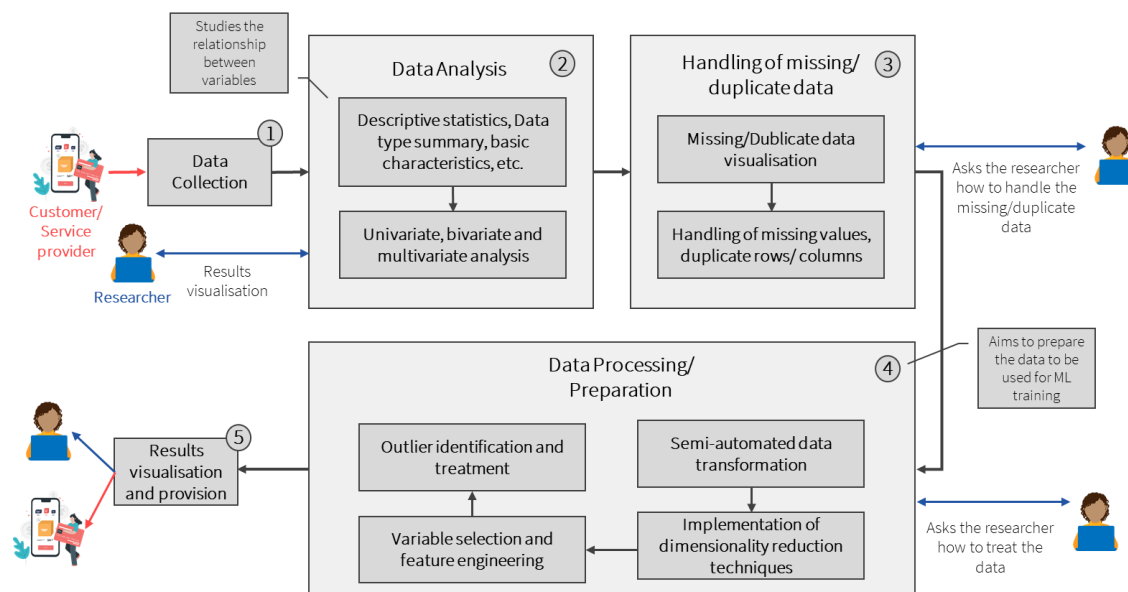


Figure 23: EDAE's High Level Architecture

5.1.1.2 Inputs, Outputs, and Interactions with Other Components

EDAE will be receiving raw data from the end-users according to the requirements set per use case. For instance, for the **Smart Hospitality** (P1) pilot use case, some example inputs could be:

- ▶ **Customer Purchasing Patterns Data:** such as purchase timestamp, purchase amount, purchased items/services, payment method, booking channel, loyalty status, customer segment, stay duration, feedback/ratings, additional services.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment	Page:	81 of 237
Reference:	D2.3	Dissemination:	PU
	Version:	2.0	Status: Final

- ▶ Customer Traits and Preferences Data: such as customer id, age, gender, nationality, travel purpose, interests/hobbies, dietary preferences, room preferences, activity preferences, booking frequency, booking lead time, average spending, travel companion.

Similarly, some indicative examples of input for the **Smart Manufacturing** (P3) pilot use case would be:

- ▶ Print Job Data: such as timestamp, job ID, printer ID, print material (i.e. type of material used in the print job, e.g. ABS, PLA, Nylon), print time, print quality (a set of quality metrics).
- ▶ Customer Data: such as customer ID, order ID, order quantity, order status, customer feedback.
- ▶ Print Quality Data: such as timestamp, quality issue type, severity, location on the printed object where the quality issue occurred, etc.

For the **Retail** (P6) pilot use case, some example inputs could be:

- ▶ Personal data: a) PII: Customers' first and last names, address, credit card details, b) non-PII: cookies, device ID, general location
- ▶ Behavioural data: transactional data, user interaction with various store functionalities, device information
- ▶ Engagement data: user interactions with the store, browsing the offered products, likes on social channels.
- ▶ Customer feedback data: customer reviews and opinions about the store and products, results of various types of surveys.

Strict requirements have not been set with regards to the supported file types for data inputs; this will depend on the end users' needs. Communication protocols and APIs will be defined at a later stage, once the position of EDAE within the architecture is finalized.

As output, EDAE will: 1) visualise the exploratory analysis results (or send the data to T6.4 to be visualised there in case we go for a single UI) and 2) provide pre-processed data, suitable for training purposes.

Anticipated Interactions with:

- ▶ End users,
- ▶ The UI interface implemented in T6.3 of the project,
- ▶ PEC+PAT (UoG) regarding the privacy assurance during the actions involved in gathering, sharing and storing of data,
- ▶ UMu regarding privacy assurance,
- ▶ Any partner in need of in-depth data-based investigation.

5.1.1.3 Significance and Added Value for TANGO

EDA, which will be applied in the context of the TANGO project, constitutes an operative approach to data analysis, aiming to improve the understanding and accessibility of results. In more detail, EDA builds upon the soundness of statistical models and hypothesis formulation, which are also used in the classical approach, to reveal hidden and unknown information from data in a form that enables analysts to obtain an immediate, direct and easy-to-understand representation of it. Hence, visual graphs are usually generated when this approach is followed, so that a more direct and trustworthy interpretation of similarities, differences, trends, clusters, and correlations are obtained through a picture, rather than through a series of numbers. In reality, EDA forms an analytical framework where the visual examination of data sets, by means of statistically significant representations, plays the pivotal role to support the formulation of hypotheses that could be tested on new data sets. The ability of comparison

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	82 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

between two concepts, for instance the dynamic experimentation on data (e.g. evaluating the results on different subsets of a same data set, under different pre-processing conditions), along with the exhaustive visualisation capabilities, empower researchers to identify outliers, trends and patterns in data, upon which new theories and hypotheses can be built.

5.1.2 Associated Pilot Use Cases

The relation of the component to the pilots described in D2.2 is summarized in the following table:

Table 31: Association of Pilot Use Cases with T5.1

Pilot	Association
Smart Hospitality (Pilot 1)	<p>Integration with NADIA hotel system for data preparation.</p> <p>The EDAE may provide:</p> <ul style="list-style-type: none"> ▶ Statistics on the customers' traits and preferences ▶ Associations, trends and patterns among customers' traits and preferences. <p>Hotel Managers will be able to receive statistics and analysis reports on the data collected across all hotels. Both Receptionists and Hotel Managers will be able to see patterns, trends and associations between a client's traits and preferences and those of other clients.</p> <p>Anonymisation of collected data so that specific data details (e.g. nationality, gender and booking dates) do not lead to identification of specific individuals.</p>
Smart Manufacturing (Pilot 3)	<p>EDAE Module enables technicians to automatically generate reports based on failed prints.</p> <p>The EDAE will provide statistics of raw print data (before training) as well as key indicators for the manager, including:</p> <ul style="list-style-type: none"> ▶ Outlier data (among failed prints) ▶ Trends and insights on which parts fail (e.g. associations between failed prints) ▶ 3. KPIs for printing tasks
Retail (Pilot 6)	<p>The Exploratory Data Analysis Engine (EDAE) can be used to provide a deep analysis of the sales data, product data, and customer data.</p> <ul style="list-style-type: none"> ▶ Dividing customers based on shopping frequency, amounts spent and recency to define their loyalty levels. ▶ Understanding shopping patterns and preferences between customers of different countries using different systems and structures. ▶ Exposing data to analysts of suppliers, new customer prospects that are setting up their business case, other markets, and the tourism industry. <p>EDAE can later be used by business/store managers to spot weak areas in a store in order to suggest areas that can be targeted for increased revenue.</p> <p>EDAE can be also used to provide a range of insights into customer behaviour. For example, exploring large datasets pertaining to customer purchasing patterns can indicate regional taste preferences, such as what is the most popular food in each country.</p>

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	83 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

5.1.3 Threat Landscape

Although there is no specific reference to attacks against EDA components, any type of attack that could compromise the integrity of the platform's data could be applicable for the EDAE as well. A lot will depend on the position of the EDAE in the system architecture. This means that if the EDAE is included in the core frame of TANGO's components, its vulnerability will depend at a large extent on the vulnerability of the overall platform. On the contrary, if EDAE is chosen to be a standalone component, then we might have to implement separate mitigation measures.

The threat vectors that EDAE faces concerns the integrity of:

- ▶ The data inserted into the system (mainly from the end-users)
- ▶ The data generated by the EDAE

5.1.4 Functional & Non-functional Requirements

Table 32: EDAE Component Functional Requirements (Author: EXUS)

ID	Description	Priority
FR-AI-EDAE-001	The TANGO system must divide customers into categories based on shopping frequency, amounts spent and recency and generate relevant reports.	H
FR-AI-EDAE-002	The TANGO system must detect and report shopping patterns and preferences between customers of different countries using different systems and structures.	H
FR-AI-EDAE-003	The TANGO system must expose data to analysts of suppliers, new customer prospects that are setting up their business case, other markets, and the tourism industry.	H
FR-AI-EDAE-004	The TANGO system must provide a deep analysis of the sales data, product data, and customer data.	H
FR-AI-EDAE-005	The TANGO system must spot weak areas in a store.	H
FR-AI-EDAE-006	The TANGO system should explore large datasets pertaining to customer purchasing patterns.	M
FR-AI-EDAE-007	The TANGO system must provide hotel (group) managers with statistics and analysis reports on the data collected across all hotels.	H
FR-AI-EDAE-008	The TANGO system must inform hotel (group) managers on patterns, trends and associations between a client's traits and preferences and those of other clients.	H
FR-AI-EDAE-009	The TANGO system must provide technicians with reports on the print (incl. failed) jobs to improve the printer operation and the printing quality.	H
FR-AI-EDAE-010	The TANGO system must provide managers with customers' data to get KPIs from them.	H
FR-AI-EDAE-011	The TANGO system should provide data scientists with data related to prints' quality issues.	M

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	84 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

FR-AI-EDAE-012	The TANGO system must be able to provide users with tools to perform fusion from multiple sources (possibly with different data structures/schemas) and visualize the results.	H
FR-AI-EDAE-013	The TANGO system must be able to generate notifications regarding the status of all EDAE processing tasks that will be sent to the visualization and the UI components.	H

Table 33: EDAE Component Non-Functional Requirements (Author: EXUS)

ID	Description	Priority
NFR-COMP-EDAE-001	The TANGO system should be compatible with all TANGO components that consume its data and with any user platform that feeds the EDAE.	M
NFR-PORT-EDAE-001	The TANGO EDAE component should be dockerised.	M

5.2 Energy efficient AI model training [T5.2]

5.2.1 Description & Functionality

T5.2 aims at providing support for the model generation in the TANGO project. This support has been designed around three different pillars, aiming at points such as the energy consumption optimization (with the federated learning component where distributed learning in different machines with an optimizes architecture for model training or the AutoML component that provides algorithms for optimizing the hyper parameter tuning phase, aiming at reducing the model training process and its power consumption), model tracking (by means of the MLOps component that allows to keep track of the different experiments performed in the model generation as well as the artifacts involved in that process), model optimisation (by means of the AutoML that automates that process) or data privacy (where the Federated Learning component eases the model generation in contexts where data cannot be shared). All the components here mentioned are to be presented in detail below:

- ▶ The **Federated Learning (FL)** component aims at providing the functionality of Federated Learning (FL) in TANGO. By using Federated Learning techniques, models can be generated from different datasets without the need of the sharing of these datasets. This can be quite useful in scenarios where predictions based on datasets that involve sensitive data or are linked to policies that do not allow their sharing are needed. In order to make use of this, different actors can get a ML model and train it with their data, adapting the model prediction capabilities to their dataset and also allowing sharing this new model with other partners with no need to share the data used for training. List of features:
 - Support different training algorithms
 - Allow the distributed training on models
 - Offer different model merging strategies
- ▶ The **MLOps** component will enable the management of the different ML models to be created in TANGO. This management will cover the whole lifecycle of the models, from their conception to the training, serving, and monitoring phase. It is intended to be used by ML and AI developers and AI system engineers. The specific set of features of this component would depend on the tool selected to support MLOps. So far, two open-source tool candidates (Kubeflow and MLFlow) are under inspection for their use in TANGO. The list of features listed below are common in the two cases, although the degree of coverage varies between the two tools. Some features are specific to the tool selected:

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	85 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- Support for integrated Notebooks
 - Support for model training
 - Support for model serving
 - Support for model registry and versioning
 - Support for creating end-to-end ML pipelines (workflow) for experimentation.
- The **AutoML** component aims at providing support for different partners in the automation of repetitive steps to design and train machine learning models. This component will help regarding technical issues involved in the process of setting up and designing predictive models (such as hyperparameter tuning or architecture selection in case of neural networks), aiming for an automatic optimization of the model to be created. In this regard, several potential tools (including AutoKeras, Katib, etc.) to support the implementation of this component are under study in order to select the one that better fits the needs to be addressed in TANGO. Some of the current features being considered are:
- Support for several techniques for machine learning model creation, such as hyperparameter tuning, or Neural Architecture Search (NAS).
 - Ease of use.
 - Configuration capabilities
 - Potential integration with other TANGO components

5.2.1.1 Internal Architecture

The high-level overview of potential interactions of the T5.2 components can be seen in Figure 24.

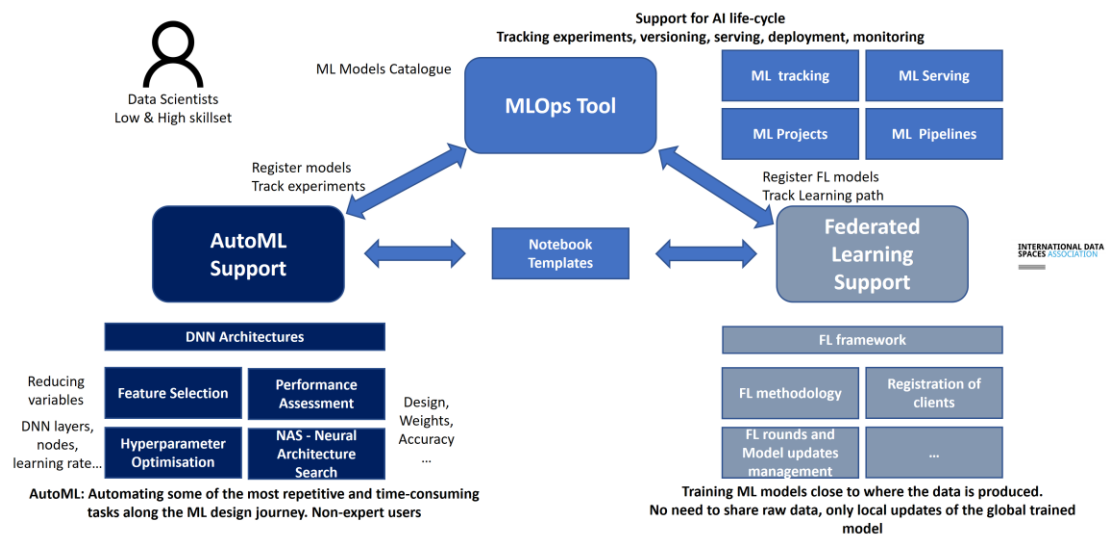


Figure 24: T5.2 Main Components and their Interactions

The figure shows that the MLOps tools play a central role (optional) to track the life cycle of ML models. There is a research activity to be carried out to enable the tracking of the FL models in the MLOps tool and its implications in terms of privacy and adequateness. AutoML models could be related to MLOps to a certain extent, at least to track the final learned models. FL and AutoML are in principle not connected.

The architecture followed in a **federated learning** process depends on the different partners involved in the use of this component. Nevertheless, the overall architecture relies on a central federated node or server (either deployed on a centralized server or directly in one of the nodes of the use cases), where the federated training and aggregation process is managed. The nodes

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment	Page:	86 of 237
Reference:	D2.3	Dissemination:	PU
	Version:	2.0	Status:
			Final

or clients are deployed on the infrastructure provided by the use cases, with connectivity to the central node and data to locally train the model. Each of these nodes will get the updated models from the federated server, train them with their local datasets and generate updates to the model, that are gathered from the central node to be aggregated. This process can be repeated until the ML model is considered final. Each user of the Federated Learning component can rely on the support provided by the TANGO decentralized architecture regarding its deployment and setup. Given the needs of the pilots involving Federated Learning, where the training of the models does not involve cross-pilot data sharing, the current view is that the Federated Server depicted in Figure 25 will be located in one of the nodes of the participants, rather than in the central TANGO architecture. This will reduce the privacy risks, the data movement, and the energy footprint.

ATOS will provide their own federated learning solution to be used in TANGO, as it is an extensible framework that will allow interoperability requirements, such as the usage of the authorisation and authentication mechanisms to be developed within TANGO.

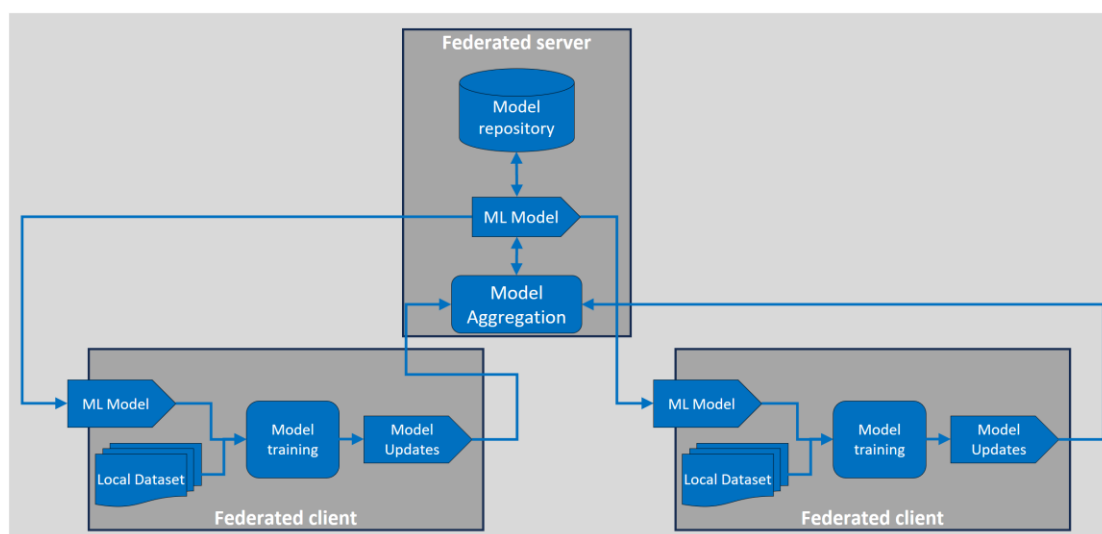


Figure 25: Federated Learning Component High-Level Description

In the case of **MLOps**, the internal description of the component depends on the tool used. Figure 26 shows the main elements of the MLOps component from a high-level, technology-agnostic perspective, depicting the different building blocks of the expected solution. MLOps will enable functionalities such as ML experiment tracking, allowing to register the model performance during the different versions produced (by training and testing) on the model building platform or development environment. Additionally, it allows not only to store models and other artifacts that might be related to the model, but also model serving (i.e. deployment for usage) to the production environment.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	87 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

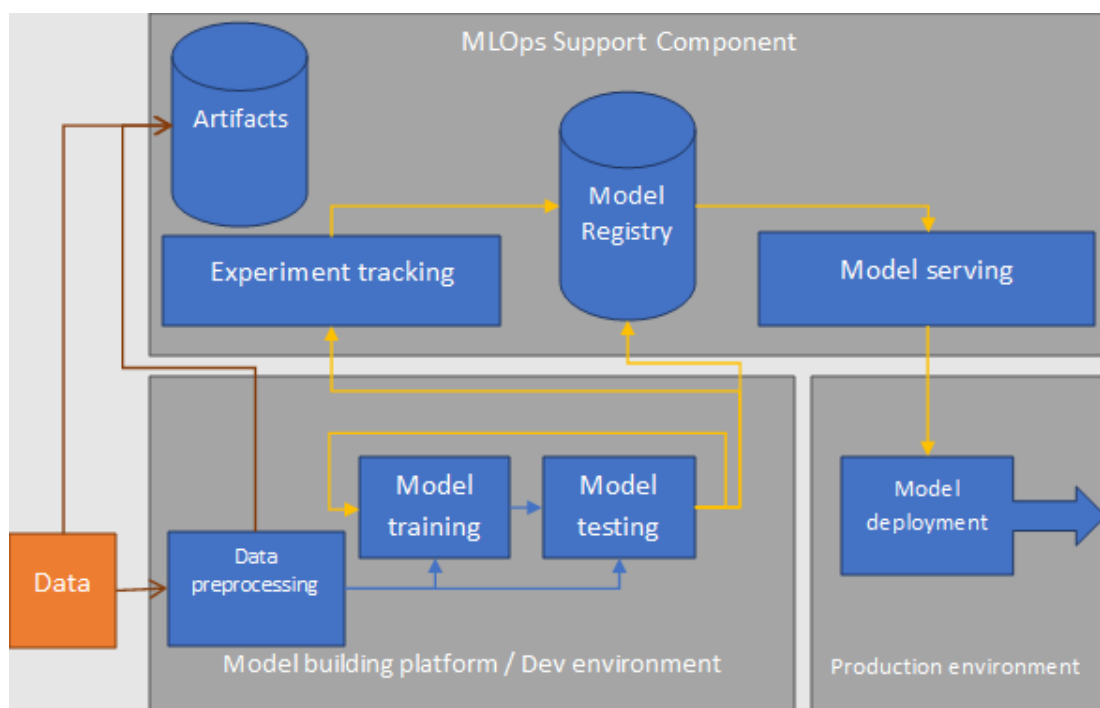


Figure 26: MLOps High-Level Depiction

As in the case of the FL component, this module could be either deployed in the central TANGO platform or in the pilot's side of platform. The former deployment would enable sharing of ML models among pilots, although it may pose issues of privacy and security if the ML models are not public. After looking at the needs of the pilots expressed in deliverable D2.2, it turns out that the use cases only require sharing ML models among participants of the pilot. This leads to the conclusion that the project does not require a central ML Model Registry for all pilots, but rather decentralised registries used for the purposes of the specific pilot use cases. After studying the functionalities offered by open source MLOps tools, such as Kubeflow and MLFlow, the decision has been made to use MLFlow. MLFlow is more lightweight and requires a lesser footprint in terms of RAM and processing power, as well as deployment requirements. Besides, nothing precludes to deploy MLFlow in the central node of the TANGO platform if required to manage future pilots' needs or even to support the ML lifecycle of specific components of the platform. Therefore, the deployment of the component is more likely to happen at the user side of the TANGO platform, as there are no requirements so far to share models among the pilots.

In the case of the **AutoML** component, although some details might vary depending on the tool chosen for its implementation, the overall architecture of this component is not complex, as shown in Figure 27. Once provided some inputs (dataset, a set of requirements and definition guidelines for the model creation), the component uses the available infrastructure for model training, generates and optimizes the model.

The functionalities and learning curve of the existing AutoML frameworks are substantial, although typically all of them offer functionalities such as feature selection, hyperparameter optimisation, or Neural Architecture Search (NAS). The decision of the tool to be used in TANGO is to a certain extent dependent on the technology used in the MLOps component. As the decision in the case of MLOps has been MLFlow instead of Kubeflow, the possibility of using Kubeflow Ktub for AutoML is less likely, although Ktub could potentially be used standalone. At the current state of the project, we are performing experiments with several tools such as Auto-Keras, H2O, or MLJAR among others.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	88 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

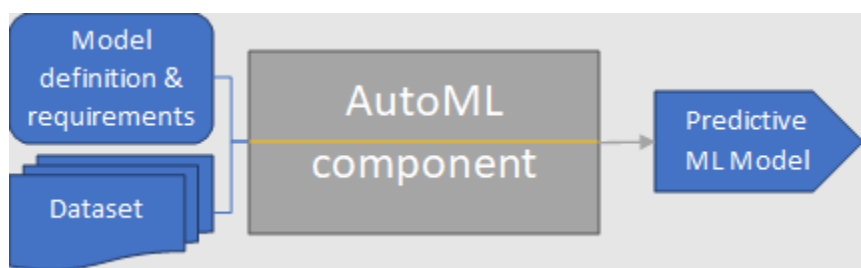


Figure 27: AutoML Component Simplified Workflow

The AutoML component could work either standalone in the user infrastructure or be deployed in the central TANGO platform. As in the previous components of these task, as the pilots don't require sharing ML models among them, the decision is that the AutoML tool will be deployed at the client side. This is also more flexible in terms of energy footprint, as the pilots would only allocate hardware resources for AutoML when required, minimising the needs of infrastructure for training in the central TANGO platform.

5.2.1.2 Inputs, Outputs, and Interactions with Other Components

The inputs of the components are the following:

- ▶ The **FL** component needs as an input the model definition to be trained, which can be either from those offered by the component (mainly Deep Neural Networks) or from models initially provided by the user. Additionally, the convergence strategy should also be defined for each model. As the architecture is distributed, the different infrastructure units involved in the federated learning process need to be appropriately set up (client registration, providing all the definition of the endpoints to interact with for each node). Additionally, the federated nodes should provide access to training datasets as input for the component, knowing that the dataset will not be sent to the federated server.
- ▶ In the case of **MLOps**, the component provides user interfaces from the tool selected to deal with the ML models life cycle in their different steps (MLFlow comes with a dedicated GUI and API). For ML developers (data scientists, ML system engineers) it should be relatively easy to use. The user would interact via Notebooks (e.g., Jupyter Notebooks), Python code, or similar, and the basics of developing, training, and serving ML models. Inputs: ML artifacts (datasets, models, metrics, interactions with the API, etc.).
- ▶ The two main inputs of the **AutoML** component are the dataset to train the model and the configuration/model hyperparameters (e.g., kind of algorithm to use for the model generation, metrics, and values to take into account in the validation of the model, model limitations such as max number of layers for a neural network, etc.).

The main outputs provided by the components are:

- ▶ In the case of **FL**, the trained ML model, that can be then registered and shared among the participants. The datasets will not be shared among the participants in the federated learning process.
- ▶ The **MLOps** component will provide means to:
 - Perform experiments with ML Models. Perform training, hyperparameter optimization, etc. The experiments will be tracked via the GUI and via specific APIs of the tools.
 - Perform model serving in different environments. The tool will be adapted to work with a set of environments and programming languages.
 - APIs and GUIs to access to the model registry, results, etc. These APIs and GUIs will likely be extensions of the ones provided by the selected tool.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	89 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ In the case of **AutoML** the expected outcomes of this component are predictive models fit to the datasets used in their creations, following the criteria set up in the configuration / definition for the model creation. These models could be registered in the Model Registry provided by the MLOps tool for further model serving in the production environment, or handled separately by the model developer.

Although these 3 components are independent, they can interact among them and with some other components. WP5 in general can be seen as a client of the TANGO platform in the sense that the components are using the data shared within the platform as input (e.g., for training the ML models).

However, several tasks in the technical WPs and in the pilots already mention the usage of AI and ML. All the components using ML are therefore potential users of this component.

Specific interactions already detected are the following:

- ▶ Special relation with the Exploratory Data Analysis tool from T5.1 is expected because the data pre-processed there could be the one used for training purposes. However, the outputs from EDAE would be curated datasets for training (persisted in a data repository and not directly consumed from one component to the other), so no specific interface is needed with this component.
- ▶ MLOps and FL or MLOps: as some artifacts or models involved in the federated learning workflow might be managed by the MLOps component, including the registration of the final resulting ML model.
- ▶ MLOps and AutoML: Same, in the case of development models and tracking their life cycle, if the underlying framework allows interoperability.
- ▶ In other tasks (e.g., T3.2), Federated Learning has been proposed as technique. This component might be used by those tasks.
- ▶ The integration with the eXplainable Artificial Intelligence (XAI) will be studied in the scope of FL and AutoML. In the case of FL, only if this integration can support scenarios where the input dataset does not need to be shared (FL). For AutoML, we will study if the underlying framework used for this component (e.g., Ktib) allows interactions with the API of the XAI library in order to get insights of the automatically developed models resulting from this component.

Special mention should be made to the relation of these tools with the authentication and authorization mechanisms implemented in WP3/4. In principle the tools used in the underlying frameworks provide in some cases their own authentication and authorization mechanisms, but it is still unclear how (or if) these can be integrated with the ones developed in TANGO. It would depend on where the tools would be deployed in the final architecture.

5.2.1.3 Significance and Added Value for TANGO

Potential for integration at the TANGO platform:

- ▶ The three components support the ML life cycle in several aspects. In most cases, these components can be considered as clients of the data acquired enabled by the TANGO platform for data sharing enabled by the WP3/4 tools of the architecture.
- ▶ As mentioned before, the MLOps and FL components need a node to manage their functionality, either residing in the central TANGO platform or at the user side. Based on the requirements from the pilots, both the MLOps tool and the FL central node will be deployed locally at the client side of the TANGO architecture, as the pilots do not need to share ML models among them and the setting poses less privacy and security risks.
- ▶ The EDAE component from T5.1 could be used to pre-process the data to be used by the components of this task.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	90 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ The XAI component from T5.5 could be integrated with the components of this task to enable explainability of the model development.
- ▶ Experiments can be made to use the TornadoVM tool from task T5.3 to accelerate some the training of specific models that match the requirements of that tool. However, this would depend on the needs of such models in the pilots.

Added value:

- ▶ In the case of **FL**, the component adds a significant added value for TANGO, as it enables some use cases where, due to the nature of the data involved in the use case, it is impossible to perform data sharing. By means of this component, it is possible for all the involved partners in those use cases to contribute to (and benefit from) the creation of better trained ML models with no need to share sensitive data.
- ▶ The technology for **MLOps** allow data scientists and ML system engineers to get support for the ML life cycle, such as the generation and management of ML models in a standardized and traceable way. ML training and experimentation gets recorded, which also provides trustworthiness to the model.
- ▶ The **AutoML** component aims at easing the model creation process for those TANGO users that are not fluent with machine learning but want to take advantage of it. These users might only need the training dataset for the model and some guidelines for the model creation.

5.2.2 Associated Pilot Use Cases

These technologies will be used in 1) the Manufacturing pilot in the FMAKE use cases for support of model training for the data scientist; 2) in the financial pilot, to train models for different banks without moving the data from the bank premises. The usage of AutoML has been also mentioned in the Retail scenario, although the decision is not final. The second manufacturing pilot (RIAS) also mentions Federated Learning as a possibility to train specific models for tracking movements, but it is more likely to finally use different techniques.

The **Federated Learning** component has been found relevant to the following pilots:

Table 34: Association of Pilot Use Cases with the Federated Learning Component of T5.2

Pilot	Association
Smart Manufacturing FMAKE (Pilot 3)	As described in the user journey of the Data Scientist of the Smart Manufacturing FMAKE use case in D2.2, the Data Scientist of the UC might need to train predictive models based on data “from different customers (stored in various locations)”. The Federated Learning Component provides the functionality of distributed training of machine learning models based on data available in different nodes, without the requirement of moving that data.
Smart Manufacturing RIAS (Pilot 3)	A part of the security systems mentioned in the RIA UC description are based on AI-powered continuous behavioural authentication, in case that model requires to be updated based on locally collected data without need of sharing that data, the usage of the Federated Learning component might be studied in case the current model is supported by the proposed component and there are resources for the model training (to be further investigated and finalized in D2.4).
Banking (Pilot 4)	As mentioned in the banking use case definition from D2.2, machine learning can be the base for a solution capable to analyse the large amounts of data depicted in the three scenarios proposed for this use case. Data sharing might offer a great opportunity for the creation of robust predictive models, but due to the nature of the data in this use case, data sharing is sometimes not possible. In this case, the inclusion

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	91 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

	of a federated learning component as the one here presented might facilitate the training of predictive models in a de-centralized way, with no need for data sharing.
--	--

The **MLOps** component will be associated with the following pilots:

Table 35: Association of Pilot Use Cases with the MLOps Component of T5.2

Pilot	Association
Smart Manufacturing FMAKE (Pilot 3)	In the Data Scientist journey described in D2.2 for the Smart Manufacturing FMAKE scenario, the Data Scientist is responsible for the good functioning of a predictive model based on the quality data outputs. The MLOps component can provide support on the versioning of that model in order to keep track of the different stages of the evolution of that model, with information of the performance of each one of the different models trained.
Banking (Pilot 4)	Due to the large amount of data presented in the banking use case, machine learning models are proposed as a solution to provide some of the functionalities proposed. In case different versions of the models to be trained are expected to be created, in order to evaluate the impact of different parameters of the models created, and always that the requirements of the use case allow that, the MLOps component might be of interest in order to keep track of these different models and their generated results.

The **AutoML**'s association with the pilots is described in the table below:

Table 36: Association of Pilot Use Cases with the AutoML Component of T5.2

Pilot	Association
Smart Manufacturing FMAKE (Pilot 3)	As described in the TANGO Persona definition of the Data Scientist in the Smart Manufacturing FMAKE scenario in D2.2, the Data Scientists “wants to build an accurate machine learning (ML) model”. In case UC allows that (i.e. the data is available and computing resources are provided), the AutoML component can provide support in the automation of the fine tuning of the machine learning model parameters in order to reach a better performance of the ML model.
Smart Manufacturing RIAS (Pilot 3)	In case the AI-powered model used for the continuous behavioural authentication system needs to be trained or improved, the AutoML component in TANGO might be used in case the current model meets the technical requirements defined for this component (to be further investigated and finalized in D2.4).
Banking (Pilot 4)	In order to ease the predictive model creation in order to process the large amounts of data depicted in the banking use case description, the AutoML component might be of consideration for its use with this purpose. This component will ease the creation of an optimised predictive model by means of evaluating different hyperparameters automatically in order to search for the best result metrics.

5.2.3 Threat Landscape

Threats:

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	92 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Federated Learning

- ▶ R1: Disclosure of model updates among participants on the FL process might provide hints about the training data.
- ▶ R2: An external attacker accesses the FL architecture and injects data into the central node.

MLOps

- ▶ R3: Potential intrusion to the component: It would mean to have access to the models or other artifacts.
- ▶ R4: Component is disabled: Access (and use) to models cannot be performed.

AutoML

- ▶ R5: Component is disabled: Automatic creation of ML models cannot be performed.

Mitigations:

- ▶ Mitigation to R1: Usage of privacy-preserving techniques or trusted mechanisms for securing the model updates.
- ▶ Mitigation to R2: TANGO authorization mechanisms are to be used when taking part in the Federated Learning component architecture.
- ▶ Mitigation to R2, R3 & R4: Infrastructure-level user access management. This way only trusted partners can have access to perform those actions they are allowed to (allow access to given models, do not allow access to component management actions).
- ▶ Mitigation to R5: locally executable alternative frameworks for AutoML could be taken into account.

5.2.4 Functional & Non-functional Requirements

Table 37: Federated Learning Component Functional Requirements (Author: ATOS)

ID	Description	Priority
FR-AI-FL-001	The TANGO system must provide the capability to setup a data processing machine as a training node of a federated learning architecture.	H
FR-AI-FL-002	The TANGO system must provide the capability to setup a data processing machine as a main node of a federated learning architecture.	H
FR-AI-FL-003	The TANGO system should support the usage of python programming language in the federated learning process	M

Table 38: MLOps Component Functional Requirements (Author: ATOS)

ID	Description	Priority
FR-AI-MLOPS-001	The TANGO system must offer the functionality to keep track of different model versions generated during experimentations	H
FR-AI-MLOPS-002	The TANGO system should be able to allow the deployment of models that have been persisted in the MLOps component.	M

Table 39: AutoML Component Functional Requirements (Author: ATOS)

ID	Description	Priority
----	-------------	----------

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	93 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

FR-AI-AUTOML-001	The TANGO system must provide a mechanism to generate a model resultant from the automatic optimization of its parameters	H
FR-AI-AUTOML-002	The TANGO system must provide metrics that allows the end user to be aware of the evaluation carried out during the model training to decide the best model to choose.	H

Table 40: Federated Learning Component Non-Functional Requirements (Author: ATOS)

ID	Description	Priority
NFR-PER-FL-001	The TANGO system must ensure that the training nodes in federated learning offer technical resources to perform the model training	H
NFR-PER-FL-002	The TANGO system should provide the capability of defining thresholds for the model performance and training.	M
NFR-SEC-FL-001	The TANGO system must ensure connectivity between the training nodes and the master node in a federated learning scenario	H

Table 41: MLOps Component Non-Functional Requirements (Author: ATOS)

ID	Description	Priority
NFR-PORT-MLOPS-001	The TANGO system must provide the MLOps component in a virtualized way	H
NFR-UI-MLOPS-001	The TANGO system should provide an interactive UI where the MLOps component user can check the stat of all the ML experiments tracked as well as all the models or artifacts registered.	M

Table 42: AutoML Component Non-Functional Requirements (Author: ATOS)

ID	Description	Priority
NFR-PER-AUTOML-001	The TANGO system must ensure that the AutoML component is ran in devices that allow the execution of the training process of the models to be generated	H

5.3 Dynamic Intelligent Execution on Heterogeneous Systems [T5.3]

5.3.1 Description & Functionality

Tornado VM [22] is an open-source programming framework that enables hardware acceleration for managed programming languages, i.e., programming languages built on top of a managed runtime. In particular, TornadoVM exposes a hardware agnostic API in Java, which includes several annotations (e.g., @Parallel, @Reduce) that can be used by a programmer. These annotations are used as a hint to signify the TornadoVM JIT compiler about the specializations that can be applied at runtime for the generation of heterogeneous code (e.g., OpenCL, CUDA, SPIRV).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	94 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

5.3.1.1 Internal Architecture

Figure 28 presents the internal architecture of TornadoVM, which comprises of three main layers:

- ▶ The **API**: exposes a set of methods to facilitate programmers to define which code segments should be offloaded for hardware acceleration along with the corresponding inputs/output data. The definition of the code segments to be accelerated is at the granularity of a method. Thus, a user defined method is mapped to a TornadoVM task. The API enables user to chain multiple user defined methods and compose a graph of tasks, called TaskGraph.
- ▶ The **Runtime**: contains a data flow optimizer layer, which parses information from a TaskGraph and applies optimization related to data movements and any dependencies between multiple tasks. The runtime also generates the Tornado Bytecodes which are used to orchestrate the compilation and execution of the user code on a targeted hardware accelerator (e.g., a GPU).
- ▶ The **Execution Engine**: encompasses a bytecode interpreter which executes the generated bytecodes in a lazy way. One of the bytecodes corresponds to the Just-in-Time (JIT) compilation of the user-defined task, which will call the TornadoVM JIT compiler to generate an OpenCL C kernel. The TornadoVM JIT compiler currently can also emit code in the form of PTX assembly, or SPIR-V binary.

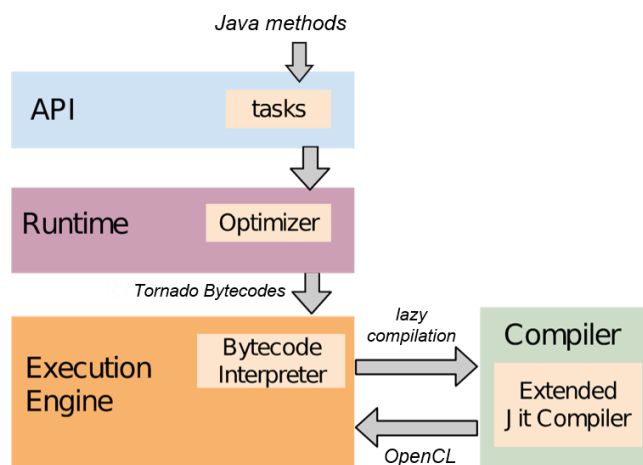


Figure 28: The TornadoVM Architecture

The flow (illustrated in Figure 28) that should be followed by a use case provider to exploit the offerings of T5.3 are:

- ▶ **Step 1:** Extend the user program (preferably in Java) to use the TornadoVM API. A guideline is sent to be published in the TANGO website as a blog.
- ▶ **Step 2:** Compile the Java program to a class file.
- ▶ **Step 3:** Invoke the tornado program and pass as input the compiled class file (Step 2).
- ▶ **Step 4:** Execute the program on a heterogeneous device (multi-core CPU, GPU) based on energy efficiency.

Features:

- ▶ Hardware acceleration on a single node. For data processing parallelization of the use cases.
- ▶ Plan execution on hardware accelerators based on energy efficiency.

The flow that describes how the software in T5.3 can be used by a TANGO data producer is shown in the figure. The input will be a Java program that uses the TornadoVM API to trigger that TornadoVM should operate in the energy efficiency mode. The output will be an execution

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	95 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

plan that will execute the user program on the most efficient device (CPU, GPU) that is available on a compute node (See image). An example of how to use the TornadoVM API is available in the TornadoVM blog series [23].

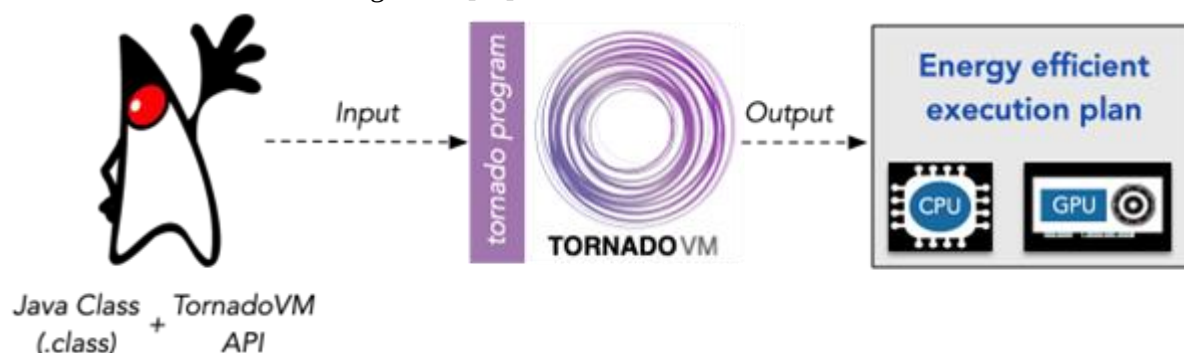


Figure 29: Overview of the Data Flow to Execute with TornadoVM

5.3.1.2 Inputs, Outputs, and Interactions with Other Components

Input: The tornado program accepts as input a Java class file (produced by the javac compiler) passed as an argument in the command line. Since, TANGO components that will utilize the offerings of T5.3 are built in Python, a Python binding will be attempted to be developed to facilitate the interfacing with Python programs.

Output: The component will execute the data processing that is defined by the user, and it will produce the result of that processing as output. The result is the data stream that is being transferred from the memory of the accelerator to the memory of the user's program via the TornadoVM API.

The component of T5.3 will be employed as an enabling mechanism that provides fast and power efficient data processing to TANGO users. The employment will be indirect to TANGO users, as the mechanism will be interacted with the EXUS's Exploratory Data Analysis Engine (EDAE) (Section 5.1) to accelerate efficiently the data processing. Thus, a Python wrapper binding of the tornado program will be developed and exposed to the EXUS programmers who will develop the functionality of the EDAE.

5.3.1.3 Significance and Added Value for TANGO

The TornadoVM technology in the context of TANGO, will allow TANGO technologies that process data to run on a heterogeneous compute node in an energy-efficient manner. Thus, it contributes to green computing and ease of programming.

5.3.2 Associated Pilot Use Cases

Task 5.3 offers a novel mechanism that can exploit hardware acceleration in an intelligent manner. This is provided via a ML-based runtime feature in TornadoVM that takes into consideration which available hardware device can be more energy efficient for the execution an application. In the context of TANGO, Task 5.3 will be provided as a side technology to the EXUS Exploratory Data Analysis Engine (EDAE - provided by Task 5.1) which aims to perform some data analytics for three use cases. Thus, Task 5.3 will be indirectly used for the hardware acceleration of the data analytics that will be performed within the EDAE engine for the Smart Hospitality, Smart Manufacturing, and Retail use cases (See Section 10.2 of D2.2, and Table 31 of this document).

Table 43: Association of Pilot Use Cases with the TornadoVM Component of T5.3

Pilot	Association				
Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	96 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Smart Hospitality (Pilot 1)	TornadoVM will be indirectly used for the hardware acceleration of the data analytics that will be performed within the EDAE engine in the context of this pilot
Smart Manufacturing (Pilot 3)	TornadoVM will be indirectly used for the hardware acceleration of the data analytics that will be performed within the EDAE engine in the context of this pilot
Retail (Pilot 6)	TornadoVM will be indirectly used for the hardware acceleration of the data analytics that will be performed within the EDAE engine in the context of this pilot

5.3.3 Threat Landscape

This is not applicable to our component (TornadoVM), as it is an internal component not exposed to the TANGO users. So, it relies on the fact that the developers involved in the project are liable for the input code.

5.3.4 Functional & Non-functional Requirements

Table 44: TornadoVM Component Functional Requirements (Author: UOM)

ID	Description	Priority
FR-AI-TVM-001	The data analytics that will be developed by the EXUS EDAE engine in Task T5.1 must not use any data dependency.	H
FR-AI-TVM-002	The data sizes that will be given for the analytics (developed by the EXUS EDAE engine in Task T5.1) must be beyond the scale of 2k elements (array length).	H

Table 45: TornadoVM Non-Functional Requirements (Author: UOM)

ID	Description	Priority
NFR-COMP-TVM-001	The EDAE engine of the TANGO system could utilize the Java API of TornadoVM in Task 5.3.	L

5.4 Privacy Threat Modelling & Identification for Trustworthy AI [T5.4]

5.4.1 Description & Functionality

T5.4 consists of two components:

- ▶ The **Privacy Enhancing Component (PEC)** which aims to provide a tool, based on the NIST Privacy impact assessment methodology, to perform privacy impact assessment.
- ▶ The **Privacy Assurance Tool (PAT)** which aims to monitor privacy related to the consent and data policy between data subjects and organisations.

Features:

PEC features:

- ▶ Identify key data actions, data types and possible threat scenarios
- ▶ Identify privacy metrics and impact values for threat scenarios
- ▶ Perform privacy impact assessment to quantify privacy risks
- ▶ Define possible countermeasures

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	97 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

PAT features:

- ▶ Collect data, consent and data policies
- ▶ Automate privacy monitoring
- ▶ GDPR compliance

5.4.1.1 Internal Architecture

PEC aims to support privacy impact assessment for data handling processes and AI/ML-driven services. The tool will define and capture components such as data actions, data types, threats, impact categories, and privacy objectives to generate privacy risk scores. The privacy risk scores will guide users to assess privacy threats, prioritise and address them. PEC will also provide a set of countermeasures to address the privacy threats.

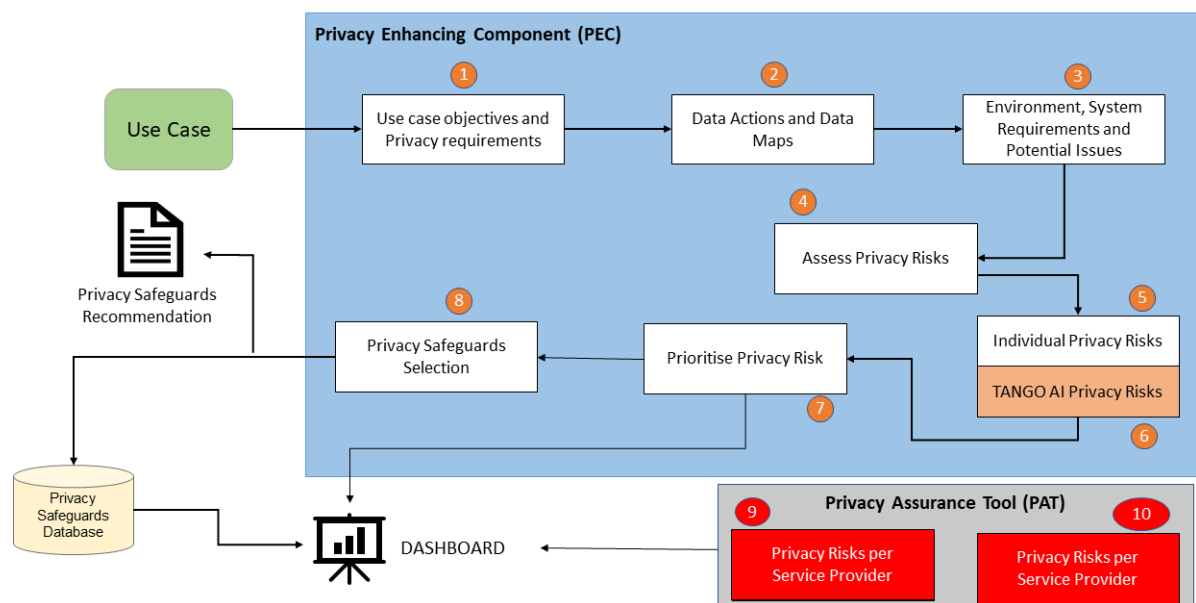


Figure 30: PEC Architecture

PAT: The architecture contains user (data subject) and organization sides. Data and consent creator and collector are used for data privacy monitoring. The user is aware about the shared data and its policy. The organization is transparent about the data purposes.

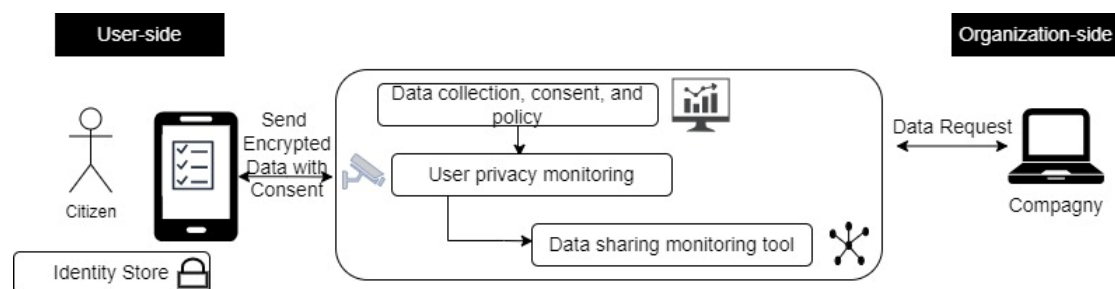


Figure 31: PAT Architecture

5.4.1.2 Inputs, Outputs, and Interactions with Other Components

PEC inputs: PEC will have a web-based interface allowing users to define data actions, critical components involved in the data action, potential attack scenarios against each critical component and how the attack could affect the NIST privacy engineering objectives

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	98 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

(Predictability, Manageability and Disassociability). Users will also be able to provide values for the likelihood of impact and loss magnitude (as defined in NIST PRAM) leading to privacy impact score for the defined threat scenario and impact pair. Furthermore, the web-based interface will also facilitate options of performing attacks on machine learning (ML) tasks such as classification or clustering to assess privacy impact from data leakage.

- ▶ Data actions
- ▶ Potential threat scenarios
- ▶ Privacy impact values

PAT inputs: PAT has its own GUI (web-based and mobile-based interfaces) from a user and organization usage perspective.

- ▶ User side:
 - Input 1: The user inserts the requested data related to a pilot use-case.
 - Input 2: The user manages their consent to enable a privacy preserving sharing of their data.
- ▶ Organization side: The organization requests for data from a user along with the associated consent and data policies.

PEC output:

- ▶ Privacy risk score
- ▶ Recommended countermeasures

PAT output:

- ▶ User side: Privacy monitoring awareness related to the shared data with organization
- ▶ Organization side: Data from users. Privacy monitoring related to the requested data from users

PEC interactions with other TANGO components:

To perform the privacy impact assessment, PEC will build upon user requirements and data handling processes defined in D2.2. This is built upon components of T3.2, especially the Usage Control component that aims to analyse privacy risks based on simulated attacks on various operations on data. Besides T3.2, PEC also aims to consider threats to data and AI identified in WP4 and WP5. Tasks such as T5.1, T5.2 and T5.5 will contribute towards identifying problematic data actions, use case driven AI/ML services and threats which will be considered to perform privacy risk assessments leading to privacy risk score and identifications of countermeasures to mitigate the privacy risks. An initial version of the proposed privacy risk assessment methodology to be used in PEC has been accepted for publication in the ARES 2023 conference.

PAT interactions with other TANGO components:

PAT will directly interact with pilots: It means that the tool will be used by the organization leading the pilots for requesting data and its consents. It will be a pilot-centred approach where the privacy assurance is managed during the data-life cycle and end-to-end data workflow: from data collection to organization and third parties for processing and learning.

5.4.1.3 Significance and Added Value for TANGO

PEC will offer a means to perform privacy impact assessment of processes that uses sensitive data. It will allow use case pilots to identify data actions that could lead to privacy leakage. PEC will also provide recommendations of countermeasures required to mitigate the privacy threats. By addressing privacy issues, organisations can assure trust and compliance with GDPR.

PAT offers:

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	99 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- ▶ Facilitation of the data sharing between a user and an organization by offering the assurance to a user to share their data.
- ▶ Privacy monitoring: The user can monitor the data sharing and its consent with an organization throughout the data lifecycle. The organization can monitor privacy risks related to the requested data and their consent.
- ▶ Transparency: The end-to-end data sharing should be transparent to the user; The user should be aware about the use of their data.
- ▶ GDPR compliance: PAT will support the organizations to fulfil their GDPR obligations.
- ▶ Security and Privacy by design: PAT should be secure by design and adhere to privacy by design requirements.

5.4.2 Associated Pilot Use Cases

The association and involvement of the PEC and PAT components to the pilot use cases is documented in the tables below (Table 46 and Table 47 respectively):

Table 46: Association of Pilot Use Cases with the PEC Component of T5.4

Pilot	Association
Autonomous Vehicles (Pilot 2)	<p>PEC will assist with privacy impact assessment for autonomous vehicles. It may support:</p> <ul style="list-style-type: none"> ▶ Identification of data actions and data elements that manage or access sensitive data. ▶ Assessment of cyber threats from privacy perspectives. This will establish how a cyber threat could lead to privacy breach and its possible impact on users. ▶ Identification of privacy threats from AI/ML-driven services. This will establish components and privacy requirements for AI/ML-driven services (eg. Object detection in autonomous vehicles) and assist with privacy impact assessment of AI/ML. ▶ Generate privacy score for each data action and recommend mitigation. ▶ Provide guidance on compliance with GDPR
Retail (Pilot 6)	<p>PEC will assist in identifying privacy leakages in AI/ML-based data processing activities for tasks such as user profiling and recommendation systems.</p> <p>It will also allow METRO to identify data processing activities involving sensitive data, perform privacy impact assessment on these data actions, get recommendations to mitigate the privacy risks and assist with GDPR compliance.</p>

Table 47: Association of Pilot Use Cases with the PAT Component of T5.4

Pilot	Association
Smart Hospitality (Pilot 1)	<p>In both smart hospitality and public administration use cases, here are the following functionalities that could be supported from the PAT component (to be further investigated and finalized in D2.4):</p> <p>In user side:</p> <ul style="list-style-type: none"> ▶ Fill-in the requested data (ex. personal data) ▶ Preferences (activities, drinks, foods, etc. in smart hospitality use case),
Public Administration (Pilot 5)	

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	100 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

	<ul style="list-style-type: none"> ▶ Data policies and privacy awareness. ▶ Monitoring privacy risks of the requested data <p>In organization side:</p> <ul style="list-style-type: none"> ▶ Request the data (ex. personal data) ▶ Data policies ▶ Data policies and privacy awareness. ▶ Monitoring privacy risks
--	---

5.4.3 Threat Landscape

PEC: Not applicable

PAT: Under investigation. Until now there is no known attacks against the technology according to recent threat intelligence reports. However, the threat landscape strategy should be periodically investigated.

5.4.4 Functional & Non-functional Requirements

Table 48: Privacy Enhancing Component Functional Requirements (Author: UOG)

ID	Description	Priority
FR-AI-PEC-001	The TANGO system must provide privacy impact assessment to identify critical data actions, threats and their possible privacy impact, and mitigation measures to be able to preserve privacy of users and assure trust on the system.	H
FR-AI-PEC-002	The TANGO system must provide a list of recommendations to address the identified privacy risks	M
FR-AI-PEC-003	The TANGO system could provide suggestions on complying with GDPR	L

Table 49: Privacy Assurance Tool Functional Requirements (Author: FUJ_LU)

ID	Description	Priority
FR-AI-PAT-001	The TANGO system must be able to collect client's personal data directly through the web. In addition, it must be able to collect the clients' preferences to provide them with personalised services.	H
FR-AI-PAT-002	The TANGO system must enable data sharing only among the relevant stakeholders/business partners.	H
FR-AI-PAT-003	The TANGO system must allow management of data based on GDPR requirements.	H
FR-AI-PAT-004	The TANGO system must provide a trusted user authentication for VISAR applicants.	H
FR-AI-PAT-005	The TANGO system must provide a trusted user authentication for VISAR employees.	H
FR-AI-PAT-006	The TANGO system must provide a secure data uploading channel.	H
FR-AI-PAT-007	The TANGO system must provide a data management approach for monitoring and managing the shared data.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	101 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Table 50: Privacy Enhancing Component Non-Functional Requirements (Author: UOG)

ID	Description	Priority
NFR-COMP-PEC-001	The TANGO system must be able to initiate request/share data among other components to support data visualisation.	H
NFR-UI-PEC-001	The TANGO system must be able to capture user input through usable interfaces to support the tasks of PEC.	H

Table 51: Privacy Assurance Tool Non-Functional Requirements (Author: UOG)

ID	Description	Priority
NFR-SEC-PAT-001	The TANGO system must provide a UI to authenticate and request data in compliance to the GDPR.	H
NFR-UI-PAT-001	The TANGO system must provide an easy-to-use UI to authenticate and share data with preferences in a trustworthy and privacy aware way.	H

5.5 X-AI For Privacy and Trust Enhancement [T5.5]

5.5.1 Description & Functionality

Explainable Artificial Intelligence (XAI) is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms. XAI is crucial for an organization in building trust and confidence when putting Artificial Intelligence (AI) models into production and it also helps the organization adopt a responsible approach to AI development.

The XAI module will be used to describe an AI model, its expected impact and potential biases. It will help characterize model accuracy, fairness, transparency and outcomes in AI-powered decision making. Apart from helping build trust in the predictions the Machine Learning (ML) models produce, it will also help debug the models in cases where bias in the decision-making process is discovered.

Main features:

- ▶ Provide Machine Learning and Deep Learning (DL) model explainability.
- ▶ Present the explainability results in an easy-to-understand visual way.

5.5.1.1 Internal Architecture

The Explainable Artificial Intelligence module for Machine Learning and Deep Learning model explainability, as most XAI libraries, will be written in Python. As there are no “one-fits-all” solutions when it comes to XAI, the explainability method that will be used will be dependent on the pilots’ models and the provision or not of the models themselves.

Depending on whether the models that need explaining are provided or not, the following methods can be followed:

- ▶ Model-specific explainability: These methods work based on the details of the specific structures of the machine learning or deep learning model which they are applied on. Model-specific XAI algorithms work by analysing the internal workings of a specific ML or DL model to provide insights into how it makes decisions. This approach, although it allows us to get a deeper understanding of the decision since we know the

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	102 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

internal structure of the model, it compromises the performance of the model because the model itself will have to be recreated.

- ▶ **Model-agnostic explainability:** These methods do not take into account the structure of the model so they can work with any machine learning model, giving more flexibility. A drawback is that these methods treat the model as a blackbox so they try to explain the output based only on the input whereas model-specific methods can explain the model's behaviour on all its layers.

The model-specific method needs for the trained model to be provided whereas the model-agnostic method does not. It should be mentioned however that even if the model is provided, it is not sure that a model specific approach could be applied, as it is difficult to decompose extremely complex models. All-in-all, despite of which explaining method will be used, the dataset that the model was trained with (training dataset) and the expected results (inferences) are a requirement to successfully produce an explainability report.

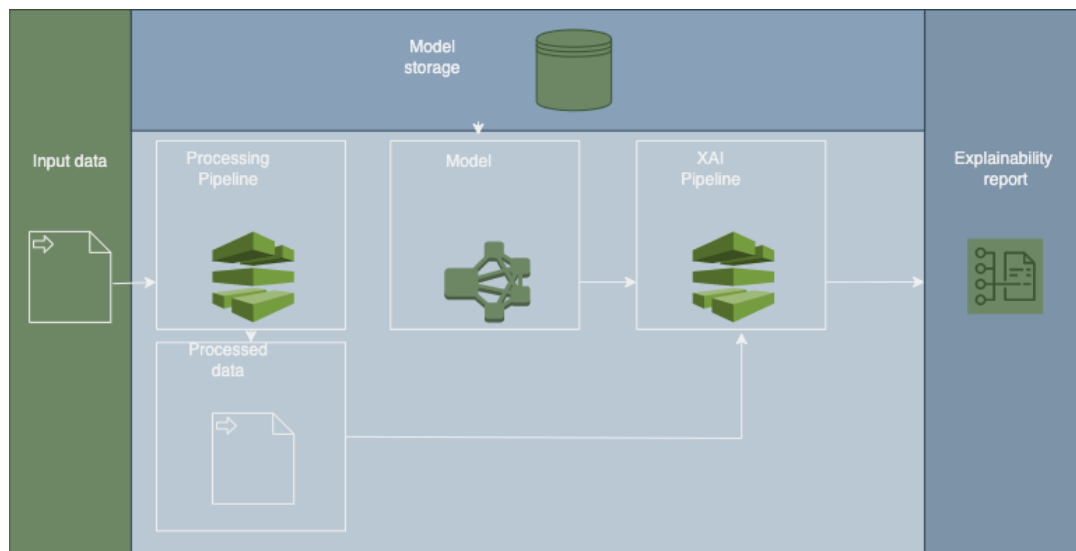


Figure 32: XAI Module Architecture

Below is a simplified explanation of the XAI module's architecture (shown in Figure 32):

- ▶ **Input data:** The input data that could be of a specific class (e.g., tabular, images, time-series)
- ▶ **Processing pipeline:** The processing pipeline that the input data is inserted to be transformed via specified functions, resulting to processed data.
- ▶ **Processed Data:** The output data of the processing pipeline is data transformed in the appropriate format required by the explainer of the XAI pipeline.
- ▶ **Model Storage:** The model storage that the necessary models are stored in.
- ▶ **Model:** This is the trained model. It is required in most of the XAI methods. In some methods (model-agnostic), other techniques will be used to compensate for the model's absence, like training a new model based on the input data.
- ▶ **XAI Pipeline:** Runs the explainer pipeline that will be used depending on the type of the data and model. For tabular data, libraries such as LIME or SHAP and for images Grad-CAM can be used.
- ▶ **The Explain-ability report:** The report that contains the results produced by the explainer of the XAI pipeline.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	103 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

5.5.1.2 Inputs, Outputs, and Interactions with Other Components

Since the training of the model will be achieved using federated learning, all data that will be needed as input reside in the device the training will take place. Those are:

- ▶ Model's training data (JSON, CSV file etc.)
- ▶ Model's inferences (JSON, CSV file etc.)
- ▶ The models that need explaining (if provided for further improving the XAI algorithm).

Output:

- ▶ Explainability report that shows which factors of the data were considered the most when the machine learning algorithm made the decision. Can be exported and sent as a JSON or CSV file (depending on the case), but it is best that this report is visualized appropriately and directly within the component. This report can also be accessed via an Application Programming Interface (API).

The XAI module's goal is to explain the models that will be trained within the TANGO platform.

The "Energy efficient AI model training" component by ATOS in T5.2 will facilitate the training of the models and the XAI module with will be used to explain the results and help with the model debugging process.

The input data that will be needed for the development of the component could be facilitated by the "Exploratory Data Analysis Engine" component in T5.1 provided by EXUS.

5.5.2 Significance and Added Value for TANGO

With so many different approaches to Artificial Intelligence, it's getting increasingly difficult for humans to determine how machines are making their decisions. With XAI we're coming closer to understanding how those decisions are made and building the trust in AI.

The pilots that are to take advantage of XAI will benefit from the explainability reports the module will produce on their existing Machine Learning models and encourage the use of AI in other aspects of their businesses, assuring that their models won't produce black boxes but rather transparent and trustworthy results.

5.5.3 Associated Pilot Use Cases

The XAI module will benefit the pilots that train their models within the TANGO platform using the "Energy efficient AI model training" components by ATOS in T5.2.

Table 52: Association of Pilot Use Cases with the XAI Component of T5.5

Pilot	Association
Smart Manufacturing FMAKE (Pilot 3)	Data scientists from the Smart Manufacturing pilot Flanders Make (FMAKE), will re-train their current quality models, and ensure that all are up to date. With the help of XAI, the debugging of the training process will be much easier and faster since the results of the model will be explained.
Banking (Pilot 4)	The financial pilot is interested in training Machine Learning models that among others will be able to detect fraudulent activities. XAI would be great at explaining why a specific activity was characterized as fraudulent and in turn detect flaws in the training process.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	104 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

5.5.4 Threat Landscape

The nature of Explainable Artificial Intelligence doesn't make it prone to adversarial attacks like model training is. If nothing else, XAI is a defence tool to help discover adversarial attacks like data poisoning during training, malicious model swapping or model evasions during the testing phase. By using XAI and explaining the results of a model's predictions, we could understand if something went wrong during the training, or if the model is compromised altogether.

5.5.5 Functional & Non-functional Requirements

Table 53: XAI Component Functional Requirements (Author: SQD)

ID	Description	Priority
FR-AI-XAI-001	The TANGO system must explain the model's predictions and describe which features had a great impact on creating the specific results.	H
FR-AI-XAI-002	The TANGO system must explain the model's predictions, giving the data scientist the opportunity to easily identify problems during the model training	H

Table 54: XAI Component Non-Functional Requirements (Author: SQD)

ID	Description	Priority
NFR-COMP-XAI-001	The TANGO XAI component must support communication over Hypertext Transfer Protocol (HTTP) with the rest of the system.	H
NFR-PER-XAI-001	The TANGO XAI component must have an optimal performance.	H
NFR-PER-XAI-002	The TANGO XAI component must have the required resources available.	H
NFR-PORT-XAI-001	The TANGO XAI component should be dockerized.	M
NFR-REL-XAI-001	The TANGO XAI component must produce reliable results.	H

5.6 Infrastructure Management based on AI [T5.6]

5.6.1 Description & Functionality

As the world faces climate change, there is a need for a shift towards greener, renewable energy resources. To realize the EUs green deal, we must make efforts in every sector to realize it. Many iterative and small improvements will make a significant impact on the EUs carbon footprint. Through TANGO we would address one of the carbon contributions: data centres. In a paper by Rong et. al [24], the authors present that data centres account for around 1.4 % of annual global electricity. A study by Oró et. Al [25] shows that the energy consumption of data centres is increasing, due to the increasing popularity of web services. EU research centre review [26] says that data centres, combined with the ICT sector, are accountable for releasing as much as 2 % of global CO2 emissions, a comparable level of pollution as of the aviation industry. By focusing on data centres and shifting even a small part of consumption made by computation, we could potentially make a significant impact on the stability of the grid and lower the carbon footprint. Furthermore, Studies [27], [28] show that the lifetime cost of electricity and cooling of a data centre will soon exceed the cost of IT equipment. Since

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	105 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Renewable energy is in some areas cheaper than fossil fuels, as source [29] states, this should reduce the operation cost of data centres as well.

Through TANGO, we would introduce **RENOPS** (Renewable Energy Forecast Production Service). Service that provides information on renewable energy potential via an easily accessible API. This information could be used to make data-driven decisions such as scheduling of energy-intensive data-centre tasks. Overall, our solution would decrease the carbon footprint of energy-intensive processing tasks. As RENOPS is provided via API its use is not isolated to cloud data centres. While its use is focused on datacentres it could be used in fog and on-premises computational environments as well. Some constraints include inability to shift time sensitive tasks, energy impact of migration of big data in case of geo-shifting and weather forecast accuracy.

RENOPS will provide features such as:

- ▶ Reduction of utility costs and carbon footprint of data centres.
- ▶ A 5-day forecast of energy availability for any location.
- ▶ Accessibility of RENOPS via API
- ▶ Delivery of information of renewable energy potential.
- ▶ Job scheduling and orchestration based on green energy availability.

Our service will provide an hourly 5-day forecast of energy availability based on geolocation. The idea is presented in the figure below (Figure 33) and goes as follows. The user initiates the request for the forecast by specifying the location of interest depicted with geographical coordinates. The service would obtain historical irradiation data and weather forecasts. The model would process the data and build the response in the form of daily renewable energy potential R_{daily} . R_{daily} is a new feature, that we introduce that will help users and orchestrators to more easily comprehend the amount of renewable energy available in contrast to only providing the W/m^2 . R_{daily} ranges between $[0, 1]$, where $R_{\text{daily}} = 0$ means that no renewable energy is available at that time for that location and $R_{\text{daily}} = 1$ would mean that 100 % of renewable energy is available of renewable type for that time and that location. This factor is the location and time-year specific since solar irradiation varies throughout the year. In practice, this means if the data centre at its peak usage of renewable energy can cover 40% of consumption with renewable energy, a potential $R_{\text{daily}} = 0.5$ will suggest that only 20% of the data centre consumption can be covered by renewable energy at a given time. To achieve this optimization, we can utilize orchestration and scheduling tools to manage the scheduled execution of these tasks automatically. Alternatively, the versatility of the RENOPS service makes it possible to provide information via various dashboards.

5.6.1.1 Internal Architecture

The internal architecture of the RENOPS service is presented in Figure 33, below.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	106 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

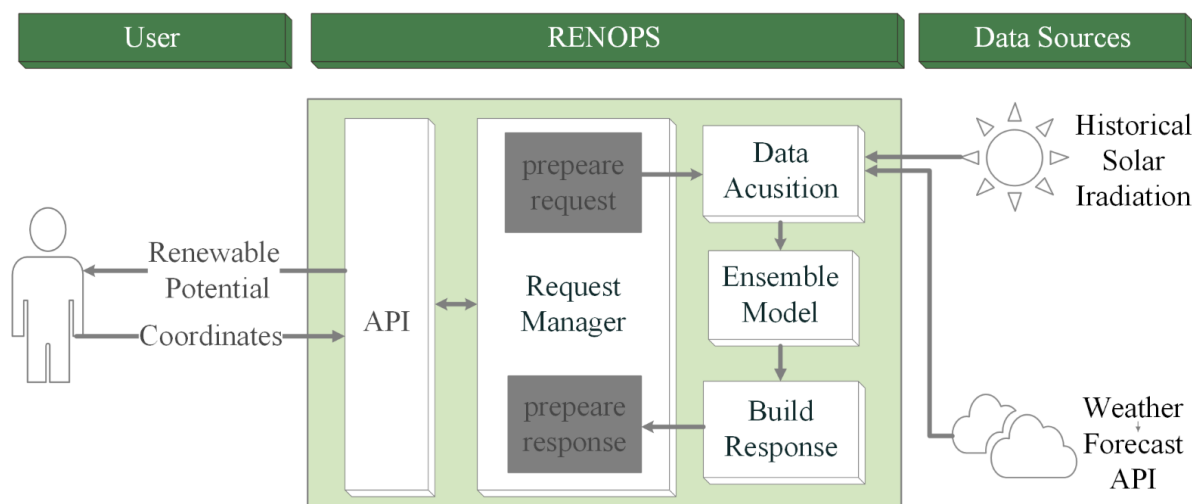


Figure 33: RENOPS Architecture and Data Flows

5.6.1.2 Inputs, Outputs, and Interactions with Other Components

Inputs to RENOPS can be split into two parts: The first input is from the user, who provides the coordinates (JSON) of the wanted forecast location. The second input into the model is fetched from various data sources such as Historical weather and forecasted future weather variables (JSON) that impact the production of renewables.

This input is fed into the ensemble model and a response in JSON format is prepared. The response is returned to the user as an output. The output of the system is not just aforementioned renewable potential but also irradiation power. Precise API routes are as follows:

- ▶ **/forecast:** Returns 5 day forecast for solar irradiation, energy, renewable factor, and cloud coverage, with hourly timestamps.
- ▶ **/forecast/power:** Returns 5 day forecast for solar irradiation with hourly timestamps.
- ▶ **/forecast/energy:** Returns 5 day forecast for solar energy with daily timestamps.
- ▶ **/forecast/renewable_potential:** Returns 5 day forecast for renewable potential with hourly timestamps.

5.6.1.3 Significance and Added Value for TANGO

As mentioned, the main goal of RENOPS is to reduce utility cost and in turn carbon footprint of computing operations. RENOPS offers many use cases, main being shifting of energy-intensive processing tasks. Examples of energy-intensive processing task can be observed in Table 55 below.

Table 55: Categorization of Energy-Intensive Processing Tasks

	Time Sensitive	Non-Time Sensitive
Shift in Time		AI model training, data analytics, backups
Shift Geographically	AI model inference	Distributed Computing

Energy intensity of big backups and big data analytics is not to be ignored, as these tasks are performed on weekly basis usually in large scales. This is one of the potential use-cases that RENOPS would address through TANGO. Second thing this component would address is making the AI/ML lifecycle less carbon intensive. For example, [30] states that training of the

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	107 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

GPT-3 language model consumed 1273 MWh of energy, which was roughly 0.012 % of Microsoft’s energy consumption in 2020. Furthermore, since renewable energy is inexpensive, and in cases of abundance even free, as we can see in ENTSO-E Transparency Platform [31], the cost of training could be greatly reduced. As models are increasing in size, complexity and in popularity these numbers are also increasing.

RENOPS could be used to shift in time (reschedule) AI model training, and since AI models are easily replicated, even geographically. In this regard this task (T5.6) could collaborate with T5.2 “Energy efficient model training”, where this solution could be used as an input to the MLOps pipeline to make data-driven decisions to schedule energy-intensive AI-model training at most optimal times.

Furthermore, RENOPS can be applied to any energy intensive process where load is controlled by an IoT device. Besides Computation, heating and cooling has a significant impact on overall consumption and in turn carbon footprint. Controlling a portion of this load could lead to same results as with datacentres’ – decrease in costs and carbon footprint.

5.6.2 Associated Pilot Use Cases

RENOPS will potentially be used as part of the smart manufacturing (FMAKE) Pilot 3 in D2.2. They utilize computer vision models as part of their workflow, meaning they must be periodically retrained using new data. RENOPS could be used to schedule the energy intensive training to find the most appropriate time slots. The same approach could be used for any pilot which makes use of periodic and energy intensive tasks such as data analytics, backups, or AI/ML re-training. Furthermore, it could also be used indirectly as part of other tasks within WP5. One example is T5.1 Exploratory Data Analysis Engine, where they process large amounts of data to analyse results. Since such calculations are usually performed periodically, RENOPS could be used to find the optimal time to execute them. A second example would be to integrate RENOPS as part of T5.2 Energy Efficient AI Model Training, where an MLOps pipeline is proposed, and RENOPS could be a feature used to make the pipelines more energy efficient.

Table 56: Association of Pilot Use Cases with the RENOPS Component of T5.6

Pilot	Association
Smart Hospitality (Pilot 1)	RENOPS could be utilised in two ways (to be further investigated and finalized in D2.4). First way would be directly via RENOPS scheduler script, that would find most optimal time to schedule backups, analytics, or AI model training. Second way would be indirectly as part of other related and energy intensive jobs as part of WP5. One possibility would be as part of T5.1 EDAE, where it could be used to shift the calculations and would be specifically beneficial in cases where calculations are periodic.
Smart Manufacturing FMAKE (Pilot 3)	To perform quality estimation of their 3D prints, FMAKE utilizes computer vision models. Larger language and computer vision models usually consume a significant amount of energy in the training and inference phase. To address the energy efficiency aspect of TANGO we plan to utilize the RENOPS system to find and schedule the training phase to the most optimal time. This should reduce the carbon footprint and utility cost of AI model training.
Banking (Pilot 4)	RENOPS could be utilised in two ways (to be further investigated and finalized in D2.4). First way would be directly via RENOPS scheduler script, that would find most optimal time to schedule backups, analytics,

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	108 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

	or AI model training. Second way would be indirectly as part T5.2 Energy efficient model training and its proposed MLops pipeline.
Retail (Pilot 6)	RENOPS could utilised in two ways (to be further investigated and finalized in D2.4). First way would be directly via RENOPS scheduler script, that would find most optimal time to schedule backups, analytics or AI model training. Second way would be indirectly as part of other related and energy intensive jobs as part of WP5. One possibility would be as part of T5.1 EDAE to find the most optimal time to run the analytics. The second possibility would be as part of T5.2 Energy efficient AI model training where RENOPS could be added as one of the features.

5.6.3 Threat Landscape

Exposing the location of users through unencrypted API responses and request is one of the threats that the RENOPS component faces. By using encryption methods used in HTTPS together with trusted CA such as Let's Encrypt [32] we would prevent from requests being unencrypted and thus preserve privacy.

5.6.4 Functional & Non-functional Requirements

Table 57: RENOPS Component Functional Requirements (Author: XLAB)

ID	Description	Priority
FR-AI-RENOPS-001	The TANGO system must provide an accurate forecast of renewable energy availability.	H
FR-AI-RENOPS-002	The TANGO RENOPS must provide the model prediction via an API.	H
FR-AI-RENOPS-003	The TANGO RENOPS system must be able to utilize information from the forecasting service and schedule energy intensive tasks in time.	H
FR-AI-RENOPS-004	The TANGO RENOPS system could utilize information from the forecasting service to geographically schedule the workload.	L

Table 58: RENOPS Component Non-Functional Requirements (Author: XLAB)

ID	Description	Priority
NFR-MAIN-RENOPS-001	The TANGO RENOPS service should be developed in CI/CD pipeline.	M
NFR-PORT-RENOPS-001	The TANGO RENOPS service should be located inside a containerized. environment.	M
NFR-REL-RENOPS-001	The TANGO RENOPS system could be managed by an orchestrator to scale the resources with demand.	L
NFR-SEC-RENOPS-001	The TANGO RENOPS system must enable secure communication and maintain integrity.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	109 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

6 Visualisation Techniques & User Interfaces

6.1 User Interfaces [T6.3]

6.1.1 Description & Functionality

This task implements the UI, User Interface, of the TANGO platform.

Thus, its functionality is to provide a web-based interface that incorporates all the options for the end-user and/or the administrator of the TANGO platform, to access and parametrize the various TANGO offerings, monitor their usage, manage users, and see statistics and information about their usage. Examples of these options will include user onboarding, identity verification, continuous authentication configuration, set and share preferences along different devices, manage the use of their personal information in relation to data protection regulations, etc. Furthermore, user and device management, access rights and monitoring of the components' usage will be available through the UI.

It is to be noted that this task does not implement the functionality of the above-mentioned options, but it rather provides the user interface for the user to access them, as they are implemented by the TANGO technology offerings. For example, user onboarding and the continuous behavioural authentication will be provided by QBE and UTH, while the privacy related functionality will be provided by UOG.

The purpose of this task is to bring the various technology offerings that will be implemented in TANGO, under a common UI. Moreover, under this common UI, the visualizations and the analytics that will be developed in Task T6.4 will be also provided, embedding them directly from their source.

The UI which will be developed aims to provide:

- ▶ Access to the technical offerings and their configuration
- ▶ A common UI for the TANGO offerings and admin tools
- ▶ A place for the data analytics and the visualization tasks to be embedded and be accessible to users.

6.1.1.1 Internal Architecture

This task implements the UI of TANGO. Since it is a front-end development task, there is no special “internal architecture”. Nevertheless, it will follow the Model-View-Controller (MVC) design to ensure that the web views are independent from the components generating the information. The MVC design is depicted in the figure below and it represents the interaction between its main concepts: the Model, the View, and the Controller.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	110 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

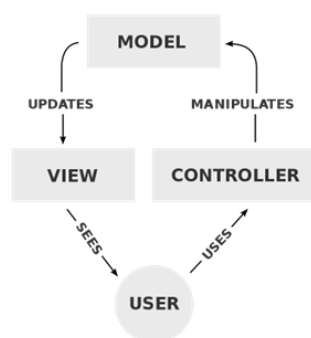


Figure 34: MVC Architecture

Image Source: [33]

In summary, the main components of the MVC design pattern are:

- ▶ **Model:** It is the central component, and it corresponds to the data-related logic; Managing data, logic, and the rules of the application.
- ▶ **View:** It is used for all the UI logic of the application, representing the information in various ways, including charts, tables, or diagrams.
- ▶ **Controller:** It acts as the interface between the model and the view. It receives user input, translates it to commands for the view and the model, and renders the final output.

This can be considered as the internal “architecture” of the UI, since it is a way of organizing the structure of the code and the data flows. If needed, a lightweight database like MongoDB will be used to temporarily store data to be visualized after receiving them from the relevant components.

6.1.1.2 Inputs, Outputs, and Interactions with Other Components

Through the UI, the TANGO offerings will present their functionality to the user. In order to achieve that, the UI will use the API output of the TANGO offerings. These APIs should be able to provide the necessary information for the user to view and edit their configuration, request user input, and present the values of attributes or the results of a processing. Thus, the “input” for the UI implementation is the various APIs of the TANGO offerings.

The output of the UI is, of course, the Web-based screens exposing the parameters and results of the technical offerings. The development of the UI will be web-based, meaning it will be implemented in a way that will be used both in the web and in the mobile, that will render it for the mobile. Since the UI will be developed separately from the back-end TANGO services, both the web and app version of the TANGO UI will use the same logic/implementation.

The UI will interact with most of the TANGO components since it will allow their usage. The final list of components that will interact with the UI will be defined at a later stage of the project implementation as it is also depending on the pilot use-cases, but the following components are foreseen to interact with the UI:

- ▶ T3.2-Trustworthy Data Sharing
- ▶ T3.3-Confidentiality and Privacy by Design
- ▶ T4.2-Seamless Onboarding for Users and Devices
- ▶ T4.3-User continuous behavioural authentication
- ▶ T4.4-Device continuous behavioural authentication
- ▶ T5.1- Exploratory Data Analysis Engine
- ▶ T5.4-Privacy Threat Modelling and Identification for Trustworthy AI

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	111 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

6.1.2 Significance and Added Value for TANGO

The User Interface is an important part of the TANGO platform. It is the point where the user will interact with the system and its services., and its services. Without the UI, the use of the TANGO offerings is not possible by the end-users, or any pilot use case.

6.1.3 Associated Pilot Use Cases

Task T6.3 is not directly related to the use cases per se. Task T6.3 implements the UI for the TANGO solution and thus, it is "horizontal" to the use cases.

The use cases will access the TANGO offerings through the user interface. UI requirements come from the technical offerings side, as it needs to support their usage and configuration. The users, being the end-users or the TANGO admins, will be able to access the TANGO offerings through the UI, view their attributes and parameters and will probably be able to configure them. The end-users are the users that are making use of the TANGO offerings, the users that are described in the pilot use cases, for example the employees of VISARIGHT, or METRO or the managers of the hotels in the hospitality use case. The TANGO admins are the users that are able to configure the TANGO offerings, monitor their usage, managing the user-access and security, viewing the related statistics. So, the end-users will be able to view the specific technology offering UI, related to the use case, or managing their TANGO account for example, while the admins will view a more generic UI, with different functionality, like user-management, connections and more technical parameters of the TANGO offerings.

So, its functional and non-functional requirements will stem out from the respective requirements of the technology offerings per use case. For this reason, we provide below the main requirements to be respected by the technology offerings in order to be able to build the UI that will support their functionality.

6.1.4 Threat Landscape

The threats against this technology are actually the threats against the UI in general and its technologies. Since this task will use Angular, the most common vulnerabilities are template injection, XSSI, authentication bypass, and more.

Apart from the technologies inherent protection mechanisms, like the ones the Angular provides, we will follow well documented and recommended techniques, that are addressing specific attacks, to secure the user interface. For now, we are not considering external technologies or software for supporting the security of the UI.

6.1.5 Functional & Non-functional Requirements

Table 59: User Interfaces Functional Requirements (Author: LST)

ID	Description	Priority
FR-S-UI-001	The TANGO UI must be able to receive messages from the system and notify the user accordingly, and vice versa.	H
FR-S-UI-002	The TANGO technical offerings must follow the OPEN API specification for providing their data and parameters to the UI	H
FR-S-UI-003	The TANGO technical offerings must use API REST back-end service that uses JSON documents for payloads. All the necessary input for the UI must use JSON documents for their payloads, in order for the UI to process them and visualize them under a common template. Specific requirements for these payloads will be drafted per technology offering, but all of them should provide an Authorization header for instance.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	112 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

FR-S-UI-004	The TANGO technology offerings must use the Swagger tool to define their interfaces, end points and the respective documentation.	H
-------------	--	---

Table 60: User Interfaces Non-Functional Requirements (Author: LST)

ID	Description	Priority
NFR-PER-UI-001	The TANGO UI must response fast to the user actions. Users should not wait long to see the options of the UI or to edit and save attributes. All their actions should be performed in a reasonable amount of time.	H
NFR-PORT-UI-001	The TANGO UI should be responsive in order to be able to efficiently work on the majority of the user devices.	M
NFR-SEC-UI-001	The TANGO UI must implement secure front-end technologies to avoid attacks. The technologies that will be used for implementing the UI must have inherent ways on protecting it from attacks. Furthermore, we should employ techniques that protect the UI from specific UI related attacks.	H
NFR-UI-UI-001	The TANGO UI must provide an easily accessible, informative, user friendly and self-explanatory way for the user to access and configure the technical offerings and their attributes.	H

6.2 Visualisation Techniques [T6.4]

6.2.1 Description & Functionality

The Visualisation and Reporting Techniques component will design and implement data visualisation interfaces to analyse big amounts of information, gathered from multiple sources and devices, and thus help the TANGO administrators to make informed, data-driven, decisions. This component is going to provide the following features:

- ▶ Visualisations for data analytics, including unstructured data from devices, structured data, processed data, and model patterns extracted from the analysis of processed data.
- ▶ Map of users' devices that will display the status of the devices, potential alerts, notifications, and recommendations.

6.2.1.1 Internal Architecture

The visualizations component will ingest the data, that could be already processed or raw data, process them or format them to the desired schema and then present them to the stakeholders using web-based modules.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	113 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

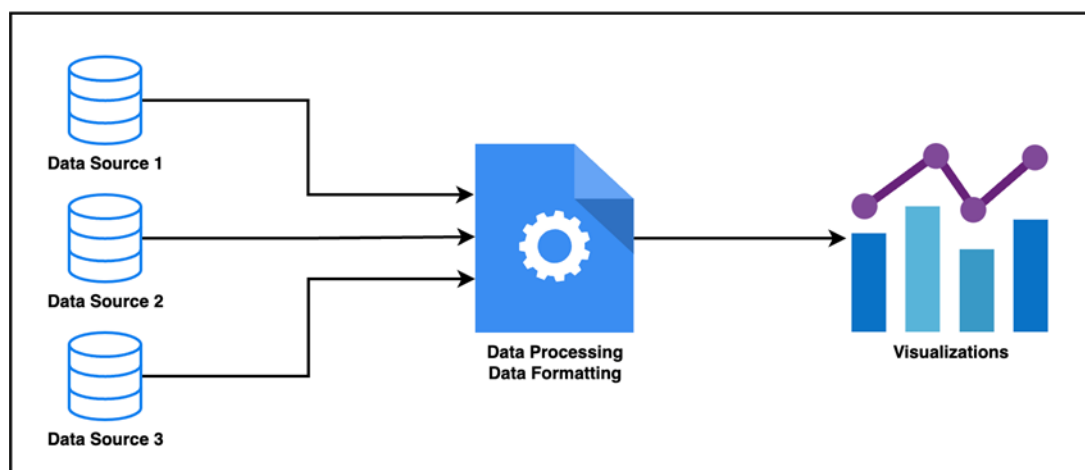


Figure 35: Visualizations Architecture

For the deployment, the visualizations will need to be integrated with the rest of the dashboards that will be built in T6.3.

6.2.1.2 Inputs, Outputs, and Interactions with Other Components

The visualization and reporting component will consume data (most likely in JSON format) that has undergone analysis from various TANGO components (like from T5.1) but also unprocessed data (e.g., from devices) that need to be processed. The output will be the visual representation of those data, as well as device mappings that will inform the administrators of their status and possible problems that might occur through alerts and notifications.

This component will create visualizations for data analytics and reports that are created from other TANGO components, along with maps of devices that different pilots own and manage.

The “Exploratory Data Analysis Engine” component in T5.1 by EXUS will execute many data analyses and will either provide the report data to be visualized or in cases of very complex reports will create the graphs so they can be integrated into the TANGO platform.

There is also high interdependency with task T6.3 that is responsible for the creation of the various TANGO platform interfaces. The graphic visualizations and device maps will need to be integrated into many parts of the platform.

6.2.1.3 Significance and Added Value for TANGO

The main purpose of the visualisation and reporting techniques is to facilitate the stakeholders to take informed decisions by providing them detailed and easy to read visual reports. Managers will be able to see complete reports and statistics about the services their organisation provides through the TANGO platform and help them improve their business. Pilots that use various tracking devices will be able to have a clear view of their status and any warnings or notifications those devices may produce. In some cases, the clients will also be able to view a summary of their experience and help them create a positive mindset for the organization.

6.2.2 Associated Pilot Use Cases

The visualization and reporting component will prove useful to many of the TANGO pilots.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	114 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Table 61: Association of Pilot Use Cases with the Visualisation and Reporting Techniques Component of T6.4

Pilot	Association
Smart Hospitality (Pilot 1)	Hotel manager and group managers from the smart hospitality pilot will be able to visually produce reports produced from data collected from customers, as well as view weekly reports about the hotels' performance and any problems that have occurred. The guests will be able to be informed about the hotel's sustainability data and environmental footprint after their stay. Furthermore, a map of the sensors installed in the rooms will be available, where the administrators will be able to check their working status and receive any notifications or alerts they may produce.
Smart Manufacturing FMAKE (Pilot 3)	In smart manufacturing, the technical experts, data scientists but also clients of Flanders Make will preview the quality reports of their corresponding printing jobs while the managers will get KPIs from data by their clients' feedback.
Smart Manufacturing RIAS (Pilot 3)	Shopfloor supervisors of RiaStone should have a clear view of all wearable devices their workers use to access the factory's controlled access areas and receive alerts if a worker enters an unauthorized area, as well as a summarized report of all unauthorized access attempts in the platform system.
Retailers (Pilot 6)	Stakeholders of the retailer pilot will view visualized reports that are produced after data analysis concerning various business matters (e.g., store and product sales) that will facilitate them to make insightful conclusions.

6.2.3 Functional & Non-functional Requirements

Table 62: Visualization and Reporting Techniques Component Functional Requirements (Author: SQD)

ID	Description	Priority
FR-S-VIS-001	The TANGO system must enable the hotel manager and the hotel group manager to visualize the reports produced by the analyzed data collected from customers.	H
FR-S-VIS-002	The TANGO system should be able to inform the clients about the hotel's sustainability data and environmental footprint in a user-friendly way.	M
FR-S-VIS-003	The TANGO system should be able to inform the hotel management and the hotel association about their sustainability data and environmental footprint.	M
FR-S-VIS-004	The TANGO system must be able to allow customers to visualize their own quality report only.	H
FR-S-VIS-005	The TANGO system must be able to allow the technical expert to visualize quality reports to help them improve the printer operation and the printing quality.	H
FR-S-VIS-006	The TANGO system must be able to produce visualizations for the reports by the analyzed customer data.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	115 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

FR-S-VIS-007	The TANGO system should be able to visually show to supervisors reports of all unauthorized attempts in the platform system.	M
FR-S-VIS-008	The TANGO system should be able to visually show to supervisors reports of all unauthorized attempts in the platform system.	M
FR-S-VIS-009	The TANGO system must allow specific users to visualize data.	H

Table 63: Visualization and Reporting Techniques Component Non-Functional Requirements
(Author: SQD)

ID	Description	Priority
NFR-COMP-VIS-001	The TANGO system must be able to initiate requests to the visualization component or send back responses to it with the appropriate data. Data sources must be accessible.	H
NFR-MAIN-VIS-001	The TANGO visualizations component must be maintainable.	H
NFR-PER-VIS-001	The TANGO visualizations component should produce the visuals in a timely manner.	M
NFR-REL-VIS-001	The TANGO visualizations component must present reliable results.	H
NFR-SEC-VIS-001	The TANGO visualizations component must be accessed only by certain users.	H
NFR-UI-VIS-001	The TANGO system visualizations component should have user friendly UI/UX	M

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	116 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

7 TANGO Reference Architecture

7.1 Introduction

7.1.1 TANGO Architecture Overview

The TANGO Reference Architecture consolidates the various heterogeneous technological assets. It includes the definition of its core components and the depiction of their relationships via an example workflow.

The architecture comprises of three layers:

- ▶ **The Trust Layer**, which includes all the services and functionalities that facilitate the data exchange and interaction between TANGO participants in a secure and trustworthy manner (i.e., Personal Credential Manager, Audit Log, Usage Control, etc.).
- ▶ **The Management Layer**, which consists of the required services and functionalities for fine-grained control, automatic deployment, and orchestration of the TANGO ecosystem (i.e., Deployment Orchestrator, Configuration Manager, etc.).
- ▶ **The User Layer**, which is comprised of the technology offerings created by the TANGO consortium and all the supporting mechanisms that enable their independent coexistence within the TANGO ecosystem (i.e., Device Continuous Behavioural Authentication, User Behaviour Exchange Module, etc.).

In the form of containerized (micro)services, the three layers are distributed across multiple runtime environments and physical locations.

The TANGO Connector is a fundamental concept of the TANGO Reference Architecture. The TANGO Connector enables the exchange of data and the hosting of data-operating services. It serves as the primary entry point to the TANGO ecosystem. Each Connector is associated with and represents a TANGO Participant, which may be a Person or User (i.e., from a pilot use-case), an IoT device, an Organization (i.e., that hosts a TANGO technology offering service), or a TANGO Administrator. The Connector hosts multiple containerized services that implement features from all three layers. TANGO Connectors can be instantiated on-premise and on user devices, or in the cloud. The TANGO Management Connector is a special type of Connector that is associated with the TANGO Administrator and is responsible for global-scope Trust and Management layer operations, including the instantiation and orchestration of all other Connectors. Blueprint configuration files provide adaptability and dynamism to the TANGO ecosystem by enabling the customization of Connectors and their hosted services, as well as the flexible orchestration of single services and service chains.

Overall, the TANGO Architecture adheres to contemporary paradigms and adopts a decentralized approach to ensuring participant trust, while providing centralized management and orchestration of the entire ecosystem.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	117 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

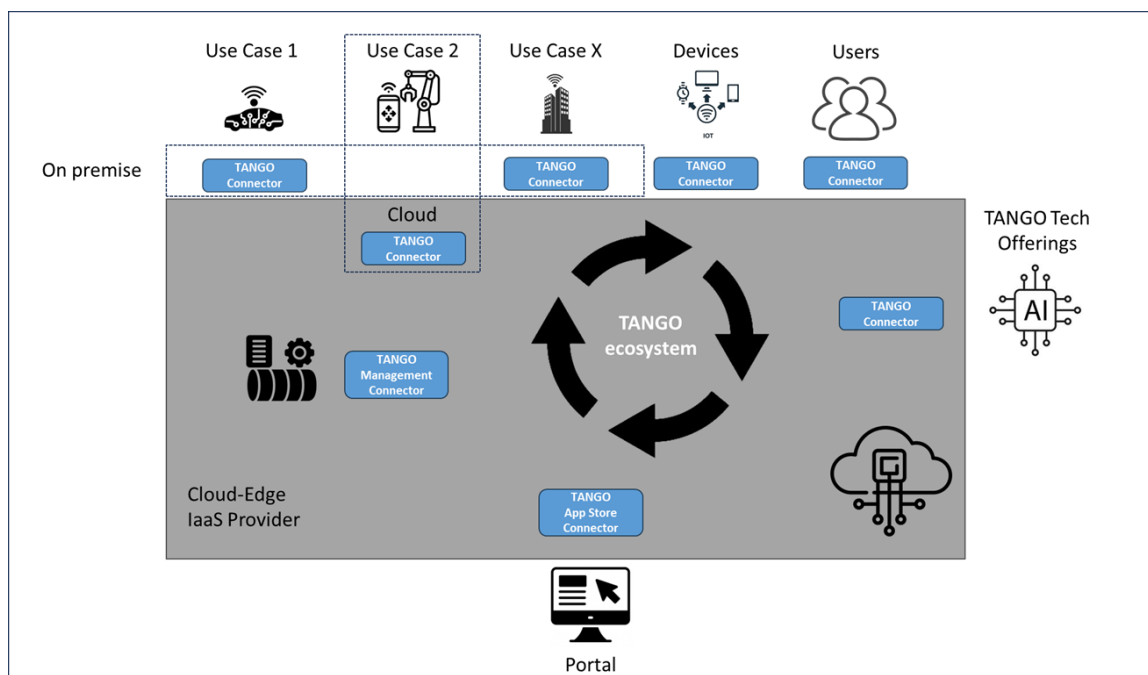


Figure 36: Overall Architecture

7.1.2 GAIA-X and IDSA compatibility

The TANGO Architecture adheres to the GAIA-X and IDSA principles by adopting several approaches and technical components from both initiatives to ensure sovereign, secure, and trusted data exchange and service provision. The TANGO Reference Architecture, like GAIA-X and IDSA, is founded on a microservices approach and incorporates the concepts of Provider and Consumer. It adopts the concept of decentralized data storage, which means that data physically remains with the data owner until being transferred outside its Connector to a trusted entity, while data and service consumers must accept the provider's usage policies, which are then enforced on data and service offerings. The Architecture promotes interoperability by permitting well-defined interaction between Connectors and is based on a decentralized identity management system for mutual authentication, eliminating the need for a central authority to ensure trust between participants.

Specifically, we have integrated into the architecture the functionality of several GAIA-X and IDSA components, and at the time of writing this deliverable we are evaluating existing open-source implementations to either tailor them to TANGO's needs or use them as-is. Such technical components are: a) the TANGO Connector, which is a refined version of the IDS Connector that offers an isolated and secure environment to host data and services, b) the GAIA-X Self-Sovereign Identity approach, c) the IDSA Clearing House for monitoring and logging data transaction in the TANGO ecosystem, d) several aspects regarding usage control from the IDSA architecture, and e) the TANGO App Store that offers similar functionalities to IDSA App Store.

However, a decision regarding the reuse of available open-source components has not yet been made, as the implementation details will be defined later during the project's development and deployment phase, following a thorough investigation of the available options. Nevertheless, Table 64 provides a preliminary list of options offered by IDSA and GAIA-X Federation Services (GXFS) that are now adopted by Eclipse foundation and are called Cross-Federation Services Components (XFSC).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment	Page:	118 of 237
Reference:	D2.3	Dissemination:	PU
	Version:	2.0	Status: Final

Table 64: IDSA and GXFS alternative implementations

TANGO Component	IDSA	GXFS
Connector	IDSA Connector	EDC Connector
Personal Credential Manager	/	PCM
Self-Encryption	/	/
Configuration Manager	ConnectorConfig	/
Audit Log	ClearingHouseService	/
Usage Control	Usage Control Service	TSA
Message Router	Apache Camel	/
Deployment Orchestrator	/	LCM Engine
Blueprint Repository	IDS-Deployment-Scenarios	/
Clearing House	Clearing House	
Verifiable Data Registry	Participant Information System ParIS	Hyperledger Indy
Trustworthiness Module	/	TRAIN
Portal	Meta Data Broker AppStore (WIP)	Portal Federated Catalogue
Self-Descriptions		Self-Description Tooling
Onboarding		Onboarding and Accreditation Workflows

7.1.3 Enabling Pilots with the TANGO Core Architecture

Several pilot use cases will be deployed and presented to demonstrate the TANGO Reference Architecture's ability to provide a flexible, user-friendly, and trustworthy platform for data sharing. Each use case requires only a subset of the TANGO ecosystem's components for successful execution. Nonetheless, the modularity and elasticity of the microservices-driven TANGO Architecture will permit the accommodation of all use case pilots concurrently, by instantiating and configuring all the necessary components either on demand or proactively. This functionality will be enabled by a "blueprint" file containing all configuration parameters of the enabling services and Connectors for each specific use case pilot. This file will be used by the system orchestrator to deploy and/or configure service chains hosted on multiple Connectors that implement all operations required for the pilots' realization.

7.2 TANGO Connector

The entire TANGO ecosystem is constituted by TANGO Connectors, each of which allowing the provision/consumption of data/services from other TANGO Connectors and as such, there is no need for a centralized point that stores data and hosts services. The Connector is essentially the point of entry for a Person, Organization, or IoT device into the secure and trustworthy TANGO ecosystem. All Connectors are accessible via the standard IP protocol and can be deployed either on-premises or in the cloud.

The TANGO Connector is an isolated and secure environment that encapsulates a number of management and trust ensuring services or technology offerings made available to the TANGO ecosystem. Each service is containerized and offers multiple data transfer and configuration endpoints. The TANGO Connector (Figure 37) is essentially a container management system (Docker, Kubernetes) that facilitates communication with other Connectors and their associated services; however, only data that is intended to be shared with other TANGO entities is made

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	119 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

public. Both internal and external communication will be based on secure protocols such as HTTPS, while the adoption of the IDS Communication Protocol is also considered.

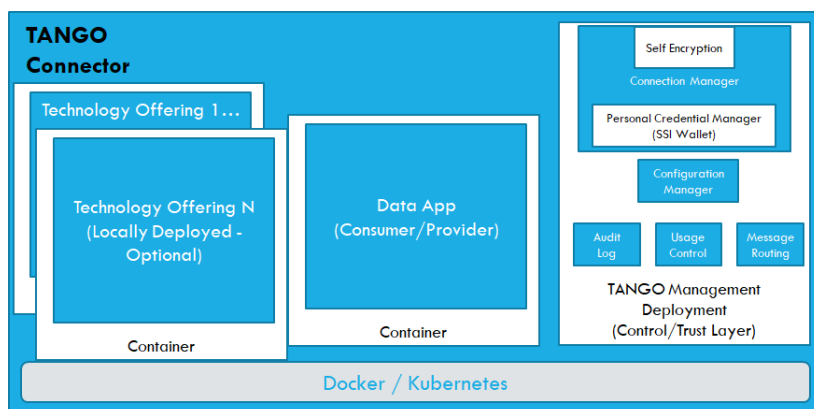


Figure 37: TANGO Connector

The Connector runs over one or multiple bare metal servers/ virtual machines or even IoT devices (although in that case it may provide limited capabilities) and hosts several containerized applications. As each Connector is associated with a TANGO technology offering or use case, at least one containerized data service consumes or provides data from across the entire TANGO ecosystem, with an associated Data App offering data processing functionalities. In addition, several containers implementing other TANGO technology offerings may be optionally downloaded from the App Store and deployed on the Connector to perform supplementary operations such as data preprocessing, hardware acceleration, etc. Moreover, each Connector contains a TANGO Management Deployment consisting of multiple containerized management services, the majority of which are client modules for various trust service components.

Typical TANGO Connector management services are described in the following subsections.

7.2.1 Connection Manager

The Connection Manager consists of two submodules and resides as a management component within each TANGO Connector. It is responsible for providing secure and trustworthy peer-to-peer communication with other Connectors across the ecosystem by authenticating with them and encrypting the data to be exchanged. These two submodules are the Personal Credential Manager (SSI Wallet), and the Self-Encryption Module.

7.2.1.1 Personal Credential Manager (SSI Wallet)

Personal Credential Manager is an SSI Mobile Wallet for Persons or an SSI Agent for devices that decentralizes the Identity Management and Authentication processes (more details can be found in Section 4.1). It stores securely the Verifiable Credentials (VCs) of the entity represented by the Connector, which are digital tokens containing identity attributes and cryptographically signed by a trusted entity, which in our case is the TANGO administrator. VCs are used as digital proof of specific attributes using Zero-Knowledge Proof techniques, ensuring trust between interacting parties and allowing selective disclosure of personal information for each transaction- the prover demonstrates the verifier the truth of the statement without revealing any information about the proof itself.

The Personal Credential Manager is responsible for the ecosystem's authentication with the other Connectors. It interacts with the corresponding modules of the other Connectors (SSI Wallets/SSI Agents) for the exchange of Decentralized Identifiers (DIDs) and with the Verifiable Credential Manager, which, potentially, consists of a blockchain/distributed ledger-

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	120 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

based Verifiable Data Registry, for the resolution of these DIDs and the retrieval of DID Documents.

Regarding this, TANGO's SSI solution plans to expand the current functionalities of the Eclipse connector (which integrates the principles of IDSA) concerning identity.

In addition, the Personal Credential Manager is tasked with managing and storing key pairs and, most importantly, the private keys of the entity associated with the Connector. Personal Credential Manager provides the interfaces for gaining access to and managing cryptography keys, as SSI and other ecosystem operations require the use of public-key cryptography to manage digital signatures and the encryption and security of sensitive data.

7.2.1.2 Self-Encryption

The self-encryption module encrypts data that will be exchanged between TANGO Connectors. Self-encryption ensures that data is encrypted prior to leaving the client's environment, so that it remains encrypted when processed or stored by other TANGO Connectors and their hosted services. This module implements the TANGO ID-based self-encryption mechanism described in Section 3.4, ensuring data ownership and multi-factor data recovery for secure data sharing among TANGO entities.

The self-encryption module communicates directly with the Personal Credential Manager, which stores the encryption/decryption keys. It is also closely related to the Attribute-Based Encryption tool, which is described in detail in Section 3.3, for encrypting the data according to an attribute-based policy defined by the data owner and decrypting them using the keys associated with the identity attributes of the Connector.

7.2.2 Configuration Manager

The Configuration Manager is responsible for receiving and applying configuration commands to the Connector, as its name implies. It configures aspects including container images, environment variables, networking, and service and endpoint discovery. It can receive configuration information as configuration commands or YAML files. YAML files are a popular, human-readable configuration format used in Kubernetes-based environments that permits the definition of structures and their relationships.

The configuration options include the following, but are not limited to:

- ▶ **Container Images:** This option specifies the container image to be used for the deployment of a particular application/service, as well as the location of the registry where the image will be located and downloaded, and potentially the authentication key to be used with the registry.
- ▶ **Environment Variables:** This option specifies the values of several environment variables to be used and accessed by the associated application/service, as well as credentials and/or configuration parameters that customize the application's operation.
- ▶ **Networking:** Configuration options used to manage networking aspects such as connectivity and communication between containers in the Connector and the external ecosystem. The Configuration Manager is responsible for establishing networking rules, such as Service IP addresses and ports, exposing services via LoadBalancers and NodePorts, and managing network policies.
- ▶ **Service Discovery:** Configuration options for intra-cluster (within the Connector) and inter-cluster (with the outside ecosystem of TANGO) DNS-based service discovery.

The Configuration Manager can apply these configuration options both as initial settings during deployment and as dynamic changes to running containers without requiring applications to be restarted. In addition, if the Connector spans multiple nodes, it is the responsibility of the Configuration Manager to synchronize their configurations in order to maintain consistency.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	121 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

7.2.3 Audit Log

The Audit Log component of the TANGO Management Deployment in the Connector is responsible for recording and optionally storing locally the details of all data transactions that occur both internally and externally with other TANGO Connectors. Each data transaction is associated with specific information, such as a unique identifier for the data, the Connector DID, the target URL of the external service with which the data is being exchanged, and data utilization and processing information.

When a data transaction concludes, it is the responsibility of the Audit Log component to package and transmit the transaction details to the Clearing House. This operation provides an additional layer of data integrity by ensuring that the Clearing House has all the necessary information to monitor and validate the data flows between the TANGO Connectors. In addition, the transaction details can be used for auditing purposes and to ensure compliance with data usage policies and GDPR regulations, thereby promoting data management transparency and accountability.

7.2.4 Usage Control

The Usage Control module is responsible for implementing and enforcing the TANGO ecosystem's data usage policies. These policies determine who may access data, under what circumstances, and for what purpose. The component is tightly integrated with the Attribute-Based policies described in Section 3.3.

While access control refers to a set of rules that define and specify who is permitted or has the rights to perform certain actions (e.g., gain access, read, write, transfer, copy) on a filesystem or more generally on the data resources of an IT system, the need to improve the control over the usage purposes of the data stored or the data flowing in a IT system, requires increased level of control capabilities and fine granularity on the data usage. To achieve this capability for enabling data usage control, it is required a data management framework that allows data producers to define and specify a set of rules that describe the usage of their data from other users (data consumers).

In the IDS Reference Model, the definition, the enforcement and the execution of specific rules, permissions and restrictions on the data usage is orchestrated by the IDS connector that performs the “Data Usage Control”. The way that IDS connector enforces data usage control policies is described in the Position Paper of members of the IDS Association named “Data Usage Control in the International Dataspaces 3.0 [34]”. The paper explains how the data usage control is implemented within IDS Connectors.

More specifically, in the IDS Reference Model the access control is based on the XACML (eXtensible Access Control Markup Language) Standard to define policies and set ABAC (Attribute-Based Access Control) rules. In the data flow model of XACML specific components (that act as control points) are responsible for the policies implementation. Namely, Policy Enforcement Point (PEP), Policy Decision Point (PDP), Policy Information Point (PIP), and Policy Administration Point (PAP). Relying on the access control model, the IDS Reference model extends it and uses it for the enforcement of usage rules by leveraging the aforementioned components to monitor, intercept and assess not only the existing but also additional information (e.g. metadata: involved stakeholders, owner, consumer, key dates, issued date, expiration date, term of the rules, for how long the rule is valid, nature and value of the rule: a legal agreement or a contract) concerning the data flows. There are three different type of conditions that refer to the data itself and are related in a timely manner with the actual time during usage (on-condition) and with the time before (pre-condition) and the time after (post condition usage). Data usage or data restriction policies follow the data using the concept of sticky policies. Those policies “stick” to the data when data is to be exchanged and follow

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	122 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

the data unless a different rule has been defined. Sticky policies define conditions and constraints about how data should be managed.

The implementation of data usage control in the TANGO architecture follows the concept of the IDS Contract and uses T3.3 exploitation asset to enforce Data Usage Control Policies by leveraging the concept of sticky policies. Sticky policies include attributes related to user roles, resource properties and specify the conditions and the constraints that must be met so that a policy can be enforced.

7.2.4.1 How Sticky Policies Enable Data Sovereignty

Data sovereignty refers to the rules that adhere to the usage and processing of data and it is linked with any action taken on the data. Moreover, it revolves around security, protection, authenticity, exchange, storage, transfer and usage of the data. Ensuring data sovereignty for the data producer necessitates the need to set data usage policies for the data flows along the data value chain path from the data producer to the data consumer. Therefore, data sovereignty requires a technical framework that defines rules on the use of data (e.g., permit data analysis and processing for a given lifetime for specific data consumer).

Sticky policies enable data sovereignty by allowing individuals or entities to maintain control over their data, even when it has been shared with others. This control is achieved thanks to the following mechanisms:

- ▶ Granular Consent Management: data owners are empowered to define specific access permissions and restrictions for their data.
- ▶ Data Portability and Consistency: data owners are able to share their data between different services while maintaining control over its use. Data owners' preferences are respected regardless of the service.
- ▶ Transparency: data owners are aware during the whole lifecycle of the data sharing process how their data is being used and by whom.

In the IDS Reference Model, the data usage control is achieved through continuous monitoring and control of the data usage policy information, which is attached to the original data. Conducting continuous data usage control on the endpoints allows for rejection or permission of processing, aggregation or forwarding to other endpoints. Thus, data producers are capable of continuously controlling their data flows without the need to control the access on specific services. Particularly, during runtime, the enforcement of the data usage control policies allows a connector to handle and process the data, or prevents the same connector to handle and process the data when this is not instructed according to the policies that follow the data.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	123 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final



Figure 38: IDS Usage Control Terms
(Image source: Position Paper for Usage Control in IDS [35])

Data usage control is achieved through a contract as shown in Figure 39, which contains apart from the actual contract data and metadata, the rules that form the usage control policy. The parties negotiating this contract conform to the settled rules which prohibit, permit and oblige for specific actions upon the data. Within the TANGO Connector, usage control enforcement is twofold: for the data provider Connector and the data consumer Connector. In the data provider Connector, the usage control restricts usage and access on data through the policy enforcement point (PEP), while in the data consumer Connector the usage control enforces the obligatory rules which in this case are handled by the policy enforcement point and the policy execution point. The interested reader could refer to the Position Paper for Usage Control in IDS [35] for more details. An exemplary paradigm of a TANGO IDS-based contract is shown in Figure 39.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	124 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

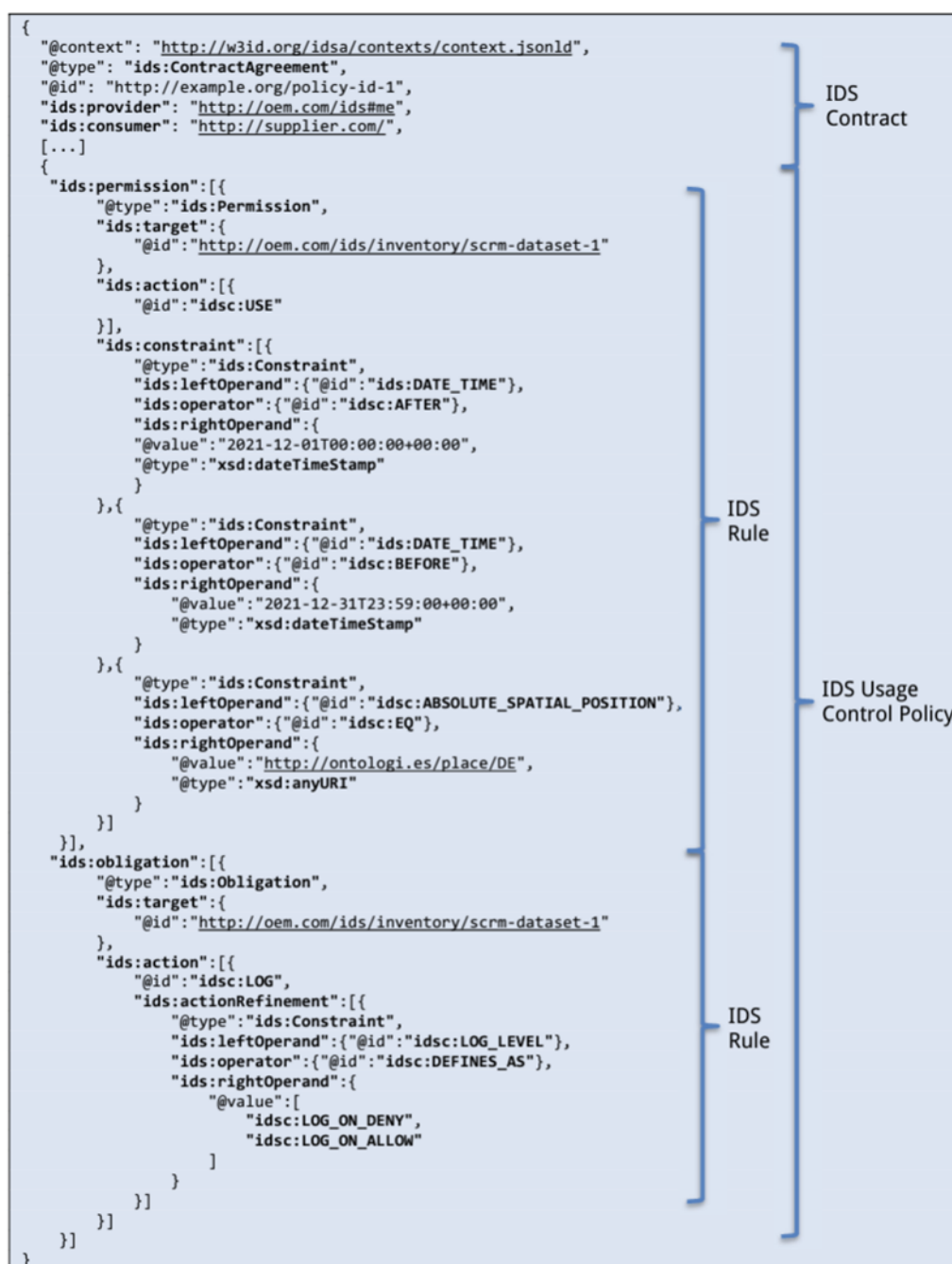


Figure 39: Example of an IDS Contract Agreement
(Image source: Position Paper for Usage Control in IDS [35])

The Usage Control module residing in the TANGO Connector implements the functionalities of PEP and PXP components and provides the configuration endpoints that enable administrators and data owners to define usage policies and translate them to contracts written in a machine-readable, technology-dependent language (e.g., MYDATA, LUCON etc.).

7.2.5 Message Router

This component serves as the data streaming and messaging layer, transferring data between TANGO Connectors. It enables the exchange of data or events between a data provider and a

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	125 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

data consuming service. This component will be implemented with popular data streaming frameworks that manage high volumes of real-time data and events, such as Apache Kafka, RabbitMQ, Apache Pulsar, etc., or with custom software tailored to the requirements of each use case with low data volumes.

The Message Router provides a configuration endpoint that permits the specification of the remote Connector’s endpoint for transferring the data and an endpoint to be used by the Usage Control module for intercepting and enforcing usage policies over the data. In addition, the Message Router will be implemented with scalability and performance in mind, capable of handling large data volumes, and with fault tolerance and resilience so that it can recover from failures and continue data transmission uninterrupted.

7.3 Blueprint

For the instantiation of every TANGO Connector, an appropriate configuration is needed in order to bring the Connector to the desired state that is required for being able to perform a specific operation in the scope of a specific scenario or a pilot use-case. Such configuration would be defined in an agreed format (i.e., most probably a YAML file) and stored into a file that will serve as the declarative definition of the Connector and its associated containerized components. It enables the definition of properties and attributes that specify how the components are created, managed and interconnected within the Connector.

As the TANGO Connectors will be based on a Container Management System such as Kubernetes, a helpful example of a configuration file like the aforementioned would be a YAML file that defines a “Deployment” in Kubernetes and which is provided below.

```

1  apiVersion: apps/v1
2  kind: Deployment
3  metadata:
4    name: my-deployment
5  spec:
6    replicas: 3
7    selector:
8      matchLabels:
9        app: my-app
10   template:
11     metadata:
12       labels:
13         app: my-app
14     spec:
15       containers:
16         - name: my-container
17           image: my-image:latest
18           ports:
19             - containerPort: 8080

```

Figure 40: YAML configuration file of a Kubernetes Deployment.

This example specifies a “Deployment” Kubernetes resource that holds three replicas of a pod that consists of a single container that is named “my-container” and is based on the image “my-image”. This “Deployment”, also, is executed as a service that exposes an endpoint at port 8080.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	126 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

There are several options that can be configured via YAML files, and besides the ones shown with this example, another useful option is the definition of a “ConfigMap”.

```

1  apiVersion: v1
2  kind: ConfigMap
3  metadata:
4    name: my-configmap
5  data:
6    my-property: my-value

```

Figure 41: ConfigMap option

A “ConfigMap” allows to define properties as key-value pairs that are later passed as environment variables within containers. Below is an extended example of the previous “Deployment” that uses the attribute “my-property” from the above “ConfigMap” in order to pass it as an environment variable to the “my-container” container.

```

1  apiVersion: apps/v1
2  kind: Deployment
3  metadata:
4    name: my-deployment
5  spec:
6    replicas: 3
7    selector:
8      matchLabels:
9        app: my-app
10   template:
11     metadata:
12       labels:
13         app: my-app
14     spec:
15       containers:
16         - name: my-container
17           image: my-image:latest
18           ports:
19             - containerPort: 8080
20           env:
21             - name: MY_PROPERTY
22               valueFrom:
23                 configMapKeyRef:
24                   name: my-configmap
25                   key: my-property

```

Figure 42: Passing environment variables to containers

Another useful option to be used in the scope of TANGO is the “imagePullSecrets” that define the encryption keys that are used to authenticate for pulling containers from a private container registry. The below code snippet extending the previous “Deployment” example illustrates the process.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	127 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

```

1  apiVersion: v1
2  kind: Secret
3  metadata:
4    name: my-secret
5  type: kubernetes.io/dockerconfigjson
6  data:
7    .dockerconfigjson: <base64-encoded-docker-config-json>
8  ---
9  apiVersion: apps/v1
10 kind: Deployment
11 metadata:
12   name: my-deployment
13 spec:
14   replicas: 3
15   selector:
16     matchLabels:
17       app: my-app
18   template:
19     metadata:
20       labels:
21         app: my-app
22     spec:
23       imagePullSecrets:
24         - name: my-secret
25       containers:
26         - name: my-container
27           image: registry.example.com/my-image:latest
28           ports:
29             - containerPort: 8080

```

Figure 43: Sample configuration file using the “ImagePullSecrets” option

The “Secret” resource contains a Docker registry credential secret, which is a base64-encoded Docker configuration JSON and the private repository from where the image of the container is pulled is indicated in the field “image” of the “containers” section.

In the scope of TANGO, we define as a “blueprint” or “recipe” a configuration file (i.e., a Kubernetes YAML file) or a collection of configuration files (i.e., a Helm chart) that are used to define and configure a complete service chain, which consists of multiple interconnected and interdependent services that collectively constitute a TANGO technology offering. The blueprint describes the desired state of services, including their configuration, environment variables, dependencies, and relationships. Blueprints provide a flexible way of deployment, configuration and management of the lifecycle of TANGO Connectors and the efficient reuse of the infrastructure in order to accommodate the needs of TANGO technology offerings and pilot use-cases.

7.4 TANGO Management Connector

The TANGO Management Connector is a special TANGO Connector that contains all the necessary services and modules for the trustworthy management of the TANGO ecosystem and its transactions. TANGO Management Connector is a container management system that hosts the TANGO Management and Trust Services and interacts with and configures the TANGO Connectors. The TANGO Management services are responsible for handling administrative and orchestration operations in the ecosystem, such as storing and managing repositories of configuration files, deploying TANGO connectors and configuring them in order to perform

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	128 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

use-case specific operations, performing access control and enforcing usage control. The TANGO Trust services establish trust relationships within the ecosystem. They handle authentication and identity management as well as evaluating the trustworthiness of entities and monitoring transactions between them.

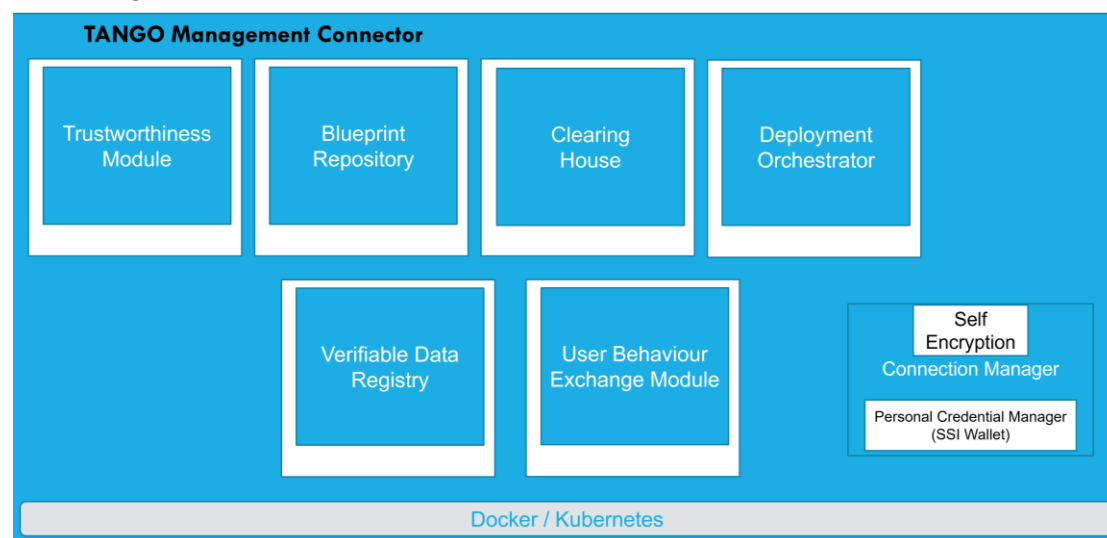


Figure 44: TANGO Management Connector

Following is a brief discussion of a subset of the TANGO Management and Trust services that are included in the TANGO Management Connector.

7.4.1 Connection Manager

It is the same module as the one described earlier for the TANGO Connector.

7.4.2 Deployment Orchestrator

The Deployment Orchestrator is envisioned as a control-plane component that ensures the efficient, consistent, and harmonious deployment of various software components and services within the data space environment. Its main objective is to bridge the sometimes-vast gap between the multiple operational environments by automating various software deployment processes.

The orchestrator will offer a robust suite of features. At its core, it will manage configurations, holding vital details for services ranging from environment variables to runtime settings. A generic lightweight resource abstraction layer will be developed, thus enabling uniform deployment and execution of various services and applications, while the execution model will be based on the adoption and adaptation of the serverless paradigm. Moreover, the orchestrator promotes automated deployments, providing tools to facilitate a seamless transition from development to production through Continuous Integration and Continuous Deployment (CI/CD) methodologies. It will also take on the role of a diligent resource manager, allocating and dynamically scaling resources like compute power, memory, and storage based on the service demands.

With every new service/application deployment request, the orchestrator will be fetching deployment configurations tailored for the target environment. It will conduct thorough dependency checks, followed by resource allocation based on specific service requirements. Services will then be deployed, adhering to a defined sequence that ensures their smooth operation and inter-service communication. Post-deployment, a continuous monitoring

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	129 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

mechanism watches over the services, ensuring optimal operations and alerting in case of discrepancies or outright failures.

TANGO Connector instantiation and/or configuration will be orchestrated by the Deployment Orchestrator module. It will be accepting a blueprint configuration file as input and instantiate and configure the service chain of a technology offering by selecting the appropriate deployment environment (e.g., on-premise, remote cloud) and interacting with the suitable configuration endpoints of the TANGO Connectors. In particular, the Deployment Orchestrator will be receiving requests from the Portal, after a TANGO user has selected a service from the existing TANGO technology offerings, to deploy, update, delete or get the status of the associated service. After retrieving and potentially updating the corresponding blueprint configuration file (YAML for Kubernetes, Ansible Playbook etc.), it performs the necessary operations on the Connectors. In the case of a service-chain deployment, the deployment is performed in phases based on the sequence defined in the blueprint for the instantiation of the individual services. In case of an existing deployment, the Orchestrator can update the offered service, for instance for providing new access to a user, by configuring its authentication service (i.e., Keycloak). In parallel, as mentioned above, the deployment orchestrator will also be receiving requests from the TANGO CI system, facilitating the continuous deployment at the development/testing and production environments. In that sense, it will be exploitable for both the developer's and end-user's side (portal). Section 8.1 provides further details in this respect.

7.4.3 Blueprint Repository

The Blueprint Repository is a centralized storage or version control system that hosts a collection of configuration files, “blueprints”, in standard formats (YAML files, Ansible Playbooks, Helm charts, Dockerfiles for docker container images creation etc.) that define and configure service chains within the context of the TANGO ecosystem. These blueprints serve as declarative definitions of TANGO Connectors and their associated containerized components, providing the necessary instructions to bring the Connectors to the desired operational state for specific scenarios or pilot use-cases. The Repository is organized for easy access and management of the blueprints and incorporates a version control system, allowing for tracking changes over time.

The Blueprint Repository enables efficient reusability of configurations and infrastructure settings, making it easier to deploy and maintain similar TANGO technology offerings or adapt existing blueprints for new use-cases. Furthermore, it incorporates access controls to ensure that only authorized users can access, modify and contribute to the blueprints and facilitates collaboration between the various service owners in order to share and contribute to the configurations and deployment plans.

Considering the aforementioned requirements, Github code hosting and version control platform [36] will be used for the blueprint repository. Selected blueprints of the repository relevant to the configuration of service chains will be accessible through the App. Store to be provided to the deployment orchestrator (section 7.4.2). Github repository will also be an integral part of the CI/CD system described briefly in section 8.1, as the main project's repository that will be supporting both daily development/testing operations and production deployments. Access control will be provided through the creation of distinct and dedicated projects within Github, where only authorized parties involved in the corresponding developments and service chain deployments will have access.

7.4.4 Clearing House

The Clearing House is a module residing in the TANGO Management Connector and that facilitates various functions related to clearing, billing, usage control and contract validation between TANGO stakeholders. The Clearing House bases all its operations on a logging service

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	130 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

that records essential information related to these operations. It integrates with the Audit Log component of the TANGO Connectors for receiving all data transaction details and supports a Usage Control Claim Validation service that uses the logged usage control data to validate usage claims on resources.

According to IDSA White Paper Specification Clearing House [37], these are the functions that the Clearing House module must fulfill during the different stages of the data-sharing lifecycle:

- Prior to data-sharing: clearing functions in terms of legal, financial and technical aspects (agreement and Usage Contract)
- During data-sharing process: monitoring and logging functions (they may be instantiated from prior stages).
- After data-sharing: settlement functions for conflict resolution in case of contract violation.

7.4.5 Verifiable Data Registry

A Verifiable Data Registry (VDR) is a system that facilitates the creation, verification, and discovery of decentralized identifiers (DIDs) and DID documents [38]. A VDR can be implemented using different technologies, such as blockchains, distributed ledgers, or web servers. Different VDRs may have different features, such as privacy, security, scalability, or interoperability.

There also exists DID methods that are much more light weight in that they do not need to rely on deploying a decentralized ledger technology among several partners with the overhead that brings on supporting a decentralized computing architecture and hence the environment.

Inline, with the TANGO approach to support a toolset of technology options that best suit the deployment scenario, TANGO will support Decentralized identifiers with or without a VDR, as described in the following sub-sections.

7.4.6 Decentralized Identifiers support without a VDR

In this project, we aim to align with the European Blockchain Services Infrastructure (EBSI) as much as possible. EBSI is a European initiative that provides a common set of standards and rules for verifiable credentials, which are a new way of expressing, exchanging, and verifying information using digital wallets and the blockchain [39]. EBSI defines two specific DID methods: one for legal entities and one for natural persons [40].

However, organisations in this project may not be entitled to use the EBSI DID method for legal entities, if not registered as legal entities by the European Blockchain Partnership (EBP) or a Member State. Therefore, we recommend using the DID:WEB method for these organisations, which allows them to have full control of their DIDs and DID documents by hosting them on their own web domains [41]. The DID:WEB method leverages the reputation and security of internet web domains, and does not require any registration or intermediaries [41], [42].

For natural persons, we recommend using the static DID:KEY method [42], which is a simple and privacy-preserving and GDPR-compliant DID method, that generates a DID and a DID document from a single cryptographic key pair. The DID:KEY method does not require any registration or resolution of DIDs or DID documents, as they are self-contained and can be stored on the wallet of the natural person. It can also be considered, the **EBSI DID Method for Natural Persons** [40], which is an **EBSI profile** for the W3C-compliant DID:KEY method. As previously indicated this does not require any registration of a DID, DID document or public keys [38], [39], and the DID and DID document are generated and stored on the wallet of the natural person [40].

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	131 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

For IoT devices, we recommend using the DID:WEB method as well, as it allows the organisations to fully control them as the controller and reset authentication keys if needed. The DID:WEB method also enables interoperability with other web-based systems and protocols that may be used by IoT devices [41].

The pros of this approach are:

- ▶ It aligns with the EBSI framework and standards as much as possible.
- ▶ It provides flexibility and autonomy for different types of entities and devices.
- ▶ It leverages existing web infrastructure and security mechanisms.
- ▶ It preserves privacy and minimises data exposure.

The cons of this approach are:

- ▶ It may introduce some complexity and overhead in managing different DID methods.
- ▶ It may require some coordination and trust among different web domains

The approach discussed above for the TANGO project, therefore promotes the use of DID Documents that do not require any Verifiable Data Registry.

7.4.7 Decentralized identifiers supported with a VDR (Fides-based approach)

As stated in section 3.1.1.4, Fides component can be used by TANGO platform in those use cases that require sharing of the regulated data. If a use case does not have strict privacy regulations, or the data pieces being shared do not constitute regulated data (for example if they are anonymised), or there is only one independent party – other methods of data sharing shall be utilized by TANGO platform, as in these use cases Fides will incur significant resource demand without adding any value.

If a use case calls for Fides usage, a VDR will be implemented as a set of Fides nodes with added capability of a standardized API interface, serving as a Universal Resolver from one side, and as a mapper between user-faced TANGO data interface and Fides internal data structures.

In a use case of decentralized sharing of regulated data there exist two conflicting requirements:

- ▶ A requirement of true decentralization, not allowing to hold all Fides nodes in the same physical location or under the same logical ownership. A significant proportion of the nodes must be separated physically and logically, to avoid having a central point of control which could become a single point of failure;
- ▶ At the same time not all independent users of TANGO platform may or are willing to afford their own installation of a Fides node on their physical or virtual (cloud) premises.

To satisfy both requirements and to allow for greater flexibility TANGO will adopt a hybrid architecture approach to decentralized storage. A decentralized Fides-based Verifiable Data Registry will comprise both private and public Fides nodes. Private nodes are the nodes installed on the premises of individual TANGO users who aim for greater control of their data, greater level of privacy and higher level of service from TANGO. These private nodes will offer a full set of capabilities to their hosts.

Public nodes on the other hand will be hosted in a public cloud environment by a TANGO operator and will offer SaaS capabilities to the TANGO users opting for this more affordable approach. These public nodes will be set up in a multi-tenant mode, offering their tenants a limited set of Fides services due to weaker privacy controls for the public nodes.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	132 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Both private and public VDR nodes will offer the same standardized TANGO API, thus allowing the owners/tenants of each node to access their own local API service. This eliminates the need for a centralized API provider.

7.4.8 Trustworthiness Module

The Trustworthiness Module is an internal module of the TANGO Management Connector that assigns scores of trustworthiness to the participating entities. It takes as input the attributes of each entity from its Verifiable Credentials and calculates a score which returns as output. This score is stored in the VC JSON-formatted file when it is calculated at registration time, or it is returned as an output following an on-demand request. It interacts with the Verifiable Data Registry to validate the Verifiable Credentials.

7.4.9 User Behaviour Exchange Module

The User Behaviour Exchange Module (UBEM) is an optional module of the TANGO Management Connector that records user actions in privacy-preserving manner, in the format of Volatile Random Number (VRN) collections. The module supports functions for providing recommendations, such as a) comparing local collection with a number of other collections, providing at least a single similarity indicator as the output, b) exchanging at least one VRN with another UBEM and c) disclosing the local collections at least partially to another UBEM. In all communications, JSON format is the default. Whenever applicable, UBEM will utilize TANGO SSI wallet as its data repository.

7.5 TANGO App Store

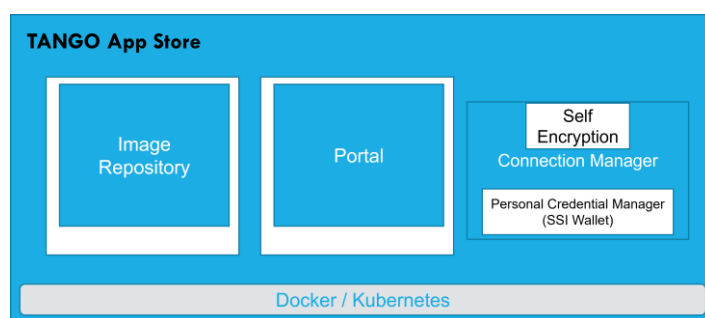


Figure 45: TANGO App Store

The TANGO App Store is a specialized TANGO Connector that allows TANGO stakeholders, users, and providers to discover or upload technology offerings and data apps. Most of TANGO technology offerings will be developed as containerized services that will consume data in order to provide some output. Indicative examples of such services can be found in Sections 5.2 and 5.5. Furthermore, some of these services are operating with a server-client architecture and as such the server part of the service will be hosted on its own Connector, whereas the client part will be deployed as a container on the TANGO Connector that consumes the service. Examples of such services are the User Continuous Behavioural Authentication (Section 4.3) and the Device Continuous Behavioural Authentication (Section 4.4), which require that there is a client-side software installed on user's device in order to monitor behaviour and a server-side performing data analytics to determine normal or abnormal activity. Another type of TANGO offering is a data app that performs data processing tasks and therefore must be deployed locally on the Connector that hosts the data source. An example of such a data app is the Exploratory Data Analysis Engine (Section 5.1).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	133 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

For all these types of TANGO technology offerings, the TANGO App Store will host the container images that will be downloaded and deployed on the Connectors, either the ones associated with a TANGO user that consumes a service or the ones hosting a service.

Overall, besides the standard management sub-modules, the TANGO App Store Connector includes two extra sub-modules: the Portal and the Image Repository.

7.5.1 Portal

The web Portal serves as an essential graphical user interface for TANGO users, allowing them to conveniently browse and inspect both technology offerings and data that their providers have uploaded onto the TANGO App Store. The TANGO technology offerings, created by the consortium members, are pushed to the Image Repository along with the corresponding metadata. This process enables the publication of technology offerings to the App Store, making them available for discovery and usage by TANGO users through the search interface offered by the Portal. This interface allows users to search for offerings based on functional or non-functional properties, potential pricing models, ratings, and other relevant criteria. The users can inspect the offerings' metadata to gain detailed information about their capabilities and functionalities and then can pull the container images from the Image Repository in order to deploy them to their TANGO Connector.

7.5.2 Image Repository

The Image Repository is a centralized storage system where container images that constitute TANGO technology offerings are stored. The images are pre-packaged, ready-to-run and can be deployed to any Connector that by default has a container runtime. The Image Repository supports versioning of the container images, so that different versions of the same technology offering can be stored, enabling easy rollbacks, upgrades, and maintenance. It integrates with container management systems like Kubernetes, enabling smooth image deployment and orchestration. Along with the container images, additional metadata is stored that provides information, such as version, authorship and usage instructions. Finally, security measures such as image signatures, checksums and digital certificates will be implemented to safeguard the integrity of container images, and authentication and authorization mechanisms will restrict access to images to specific users or groups.

For hosting and managing container images according to the aforementioned requirements, different software solutions will be examined (e.g., Docker registry [43], Jfrog artifactory [44]). The private image repository will be accessible at the end-user side through the App. Store portal to select and instantiate the selected technology offering(s)/service chain. The private image repository will also be accessible at the TANGO developer/tester side through the CI/CD system for automated software testing and deployment pipelines (described in section 8.1) to retrieve and deploy the target containers at the corresponding connector (Kubernetes) environment.

To secure the private image repository, strong authentication and access authorization measures will be put in place, through the use of security certificates and connections over HTTPS and access control to make sure that only authorized parties have access to specific software artifacts.

7.6 Example Workflow

This subsection provides an example of a TANGO ecosystem workflow. This workflow describes an example generic scenario in which a TANGO user wishes to consume a service-based technology offering. To better illustrate the interactions and dependencies between the various components of TANGO, we assume that this technology offering requires continuous

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	134 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

authentication of the user's device, and as such, it is dependent on the Device Continuous Behavioural Authentication (DCBA, T4.4) service. We illustrate the instantiation and orchestration of the service chain in order to support the deployment of the technology offering as well as the interactions and data exchanges between the TANGO Connectors.

Whenever the term user is encountered in the following text, the reader may safely assume that the term can refer also to a machine or an IoT device that interacts with the system. Moreover, each of the involved services, which are encapsulated in TANGO Connectors, can be deployed either locally or in a centralized cloud location.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	135 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

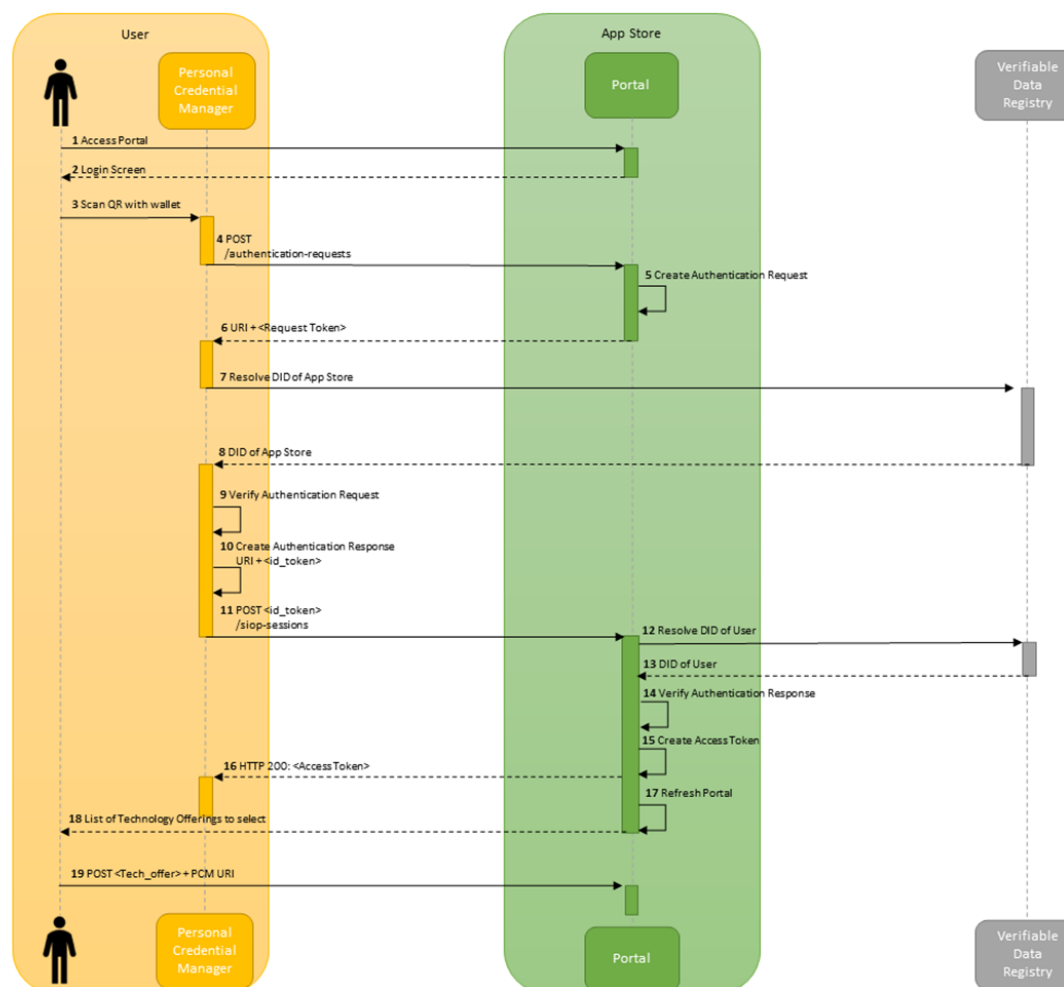


Figure 46: Example Workflow (1/3)

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	136 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

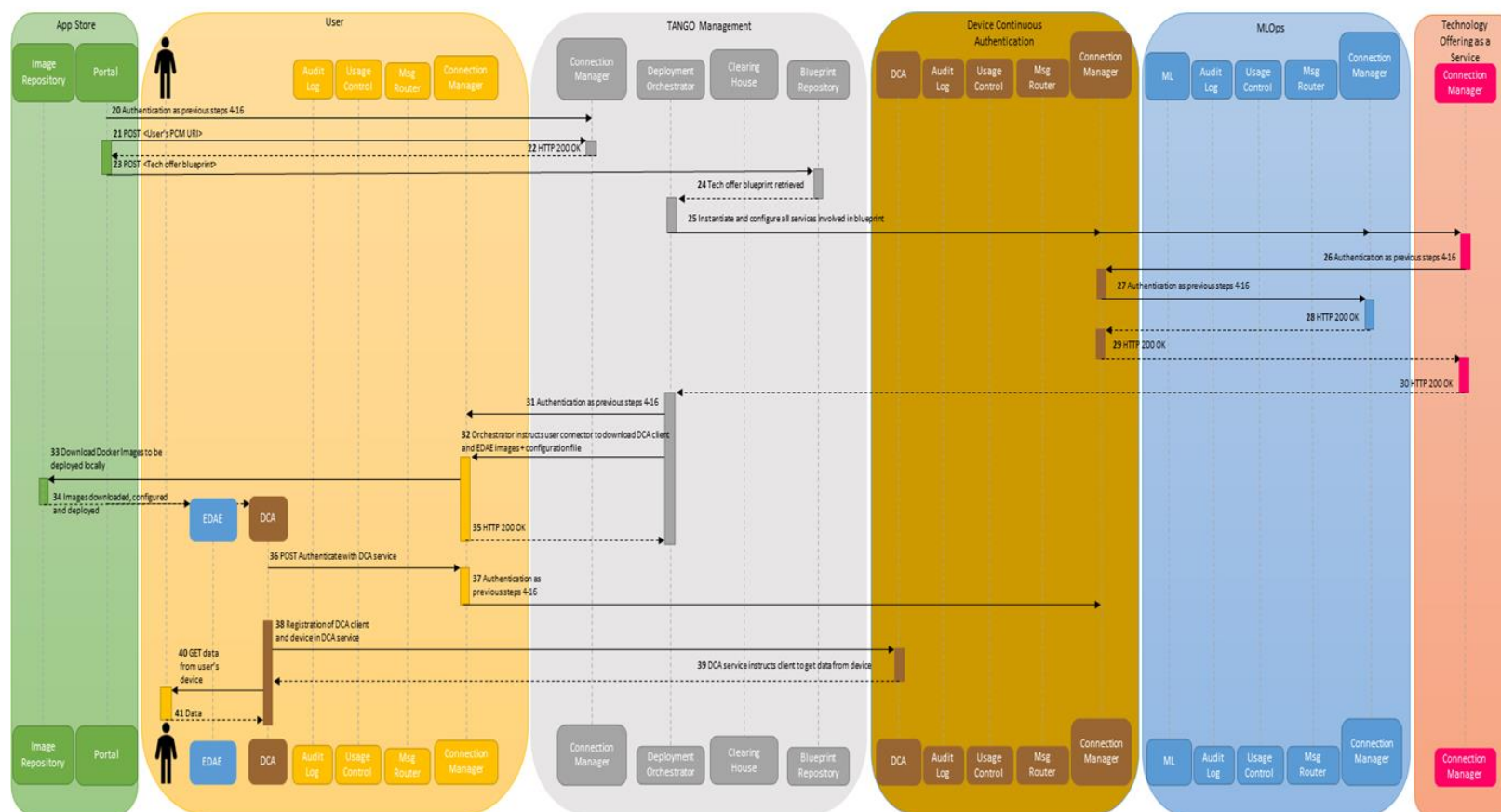


Figure 47: Example Workflow (2/3)

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	137 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

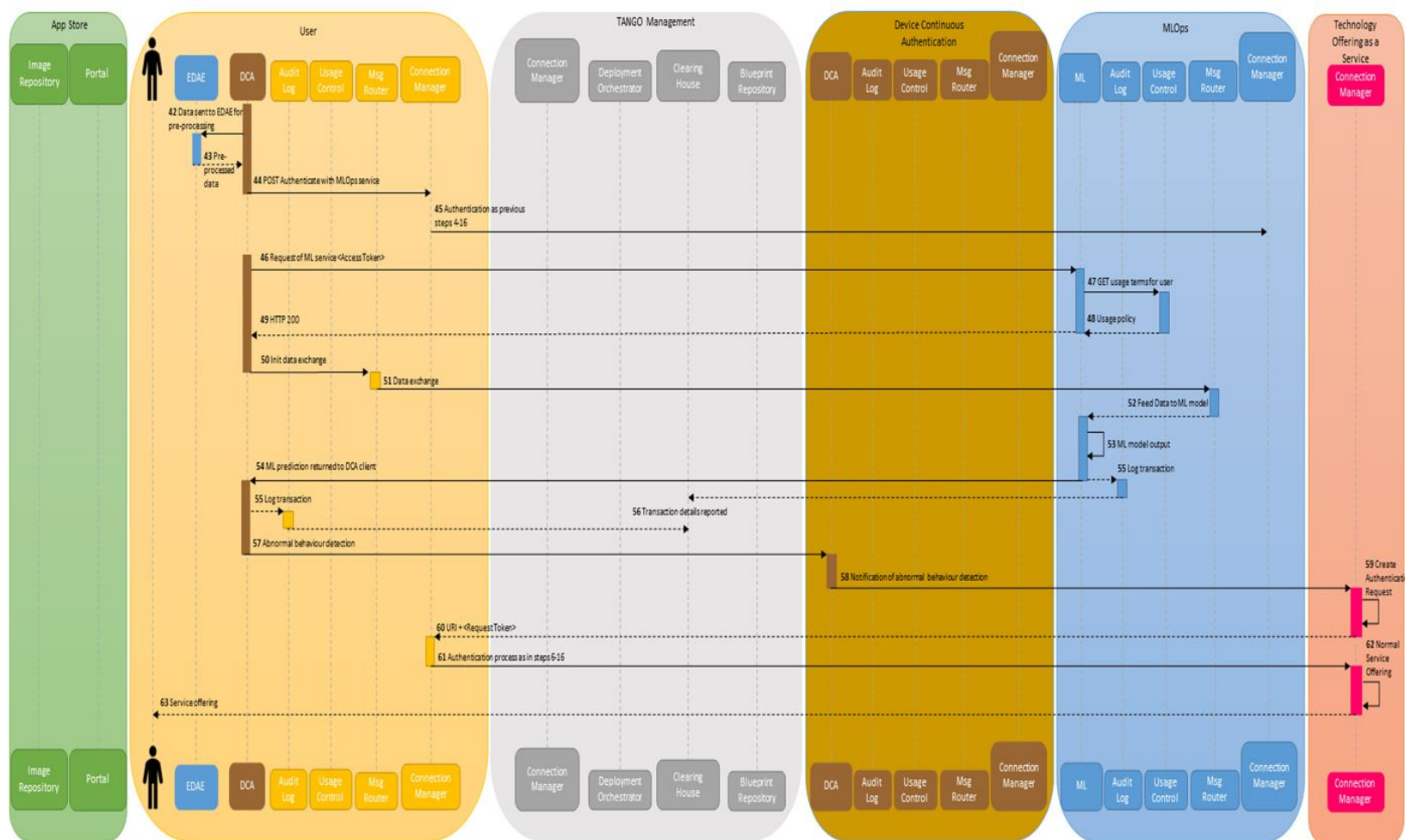


Figure 48: Example Workflow (3/3)

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	138 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

7.6.1 Steps

Figure 46, Figure 47, and Figure 48 depict all the performed steps of the workflow for selecting, deploying and instantiating a TANGO technology offering on the TANGO platform. Each numbered step is described in the following text.

1. User accesses the TANGO App Store portal in order to login and select the desired technology offering.
2. The TANGO App Store portal displays a QR that contains the URL of the /authentication-requests endpoint of the TANGO App Store Connector. In case that a device tries to authenticate with the App Store, then no QR code is generated, instead a URL endpoint is returned to the device.
3. The user scans the QR with his SSI Wallet (Personal Credential Manager). This step is omitted when a device is authenticating.
4. The Wallet calls the /authentication-requests endpoint.
5. A standard SIOP (Self-Issued OpenID Provider) flow starts, with the TANGO App Store portal having the role of Relying Party (RP) and the user the role of the OpenID Provider (OP). In this step, the App Store portal creates an SIOP Authentication Request. The Authentication Request goes back to the user through the HTTP POST request performed in the previous step, as a JWT (Json Web Token) signed by the App Store portal.
6. The Authentication Request is returned to the user's Wallet. Along with the Authentication Request token, a redirect URI is returned that indicates the endpoint where the SIOP shall deliver the authentication result.
7. In this step, the user verifies that the TANGO App Store portal is a trusted entity by resolving the DID (Decentralized Identifier) that was received as a part of the Authentication Request. For resolving the DID, the Wallet of the user sends a GET request to the endpoint of one of the trusted servers implementing the Universal Resolver (Verifiable Data Registry in this example) functionality. The Universal Resolver may include one or multiple blockchain nodes. Its role is to resolve DID Documents that contain relevant information about the entity owner of the DID. At least one of the trusted blockchain nodes must be configured in the Wallet of the user.
8. The user's Wallet receives the DID document of the TANGO App Store portal, with trusted information about the entity, including the Public Key associated with the Private Key that the App Store uses to digitally sign the tokens.
9. The user's Wallet verifies the Authentication Request, including checks on whether the TANGO App Store's DID method and key types are supported and whether the user can satisfy the requested Verifiable Credentials.
10. In this step, user's Wallet creates an Authentication Response to be posted in the redirect URI specified by TANGO App Store specified in step 6.
11. The SIOP (Personal Credential Manager) sends the Authentication Response to the endpoint specified previously, using an HTTP POST request. The Authentication Response contains an ID Token and a VP (Verifiable Presentation) token that includes the Verifiable Credentials of the user.
12. The TANGO App Store uses its own blockchain node or one of the trusted that has been configured with and that implement the Universal Resolver functionality to resolve the DID of the user, which is inside the Verifiable Credential received in the previous step.
13. The TANGO App Store receives the DID document of the user that includes the Public Key associated with the Private Key that the user used to digitally sign the Verifiable Credential. Using the Public Key, the TANGO App Store can verify the signature of the Verifiable Credential and that the user is a trusted entity in the TANGO ecosystem.

14. In this step, the TANGO App Store verifies that the Verifiable Presentation including the Verifiable Credentials of the user is sent by the user and not from a malicious agent. It does so by using the Public Key of the user.
15. Once all verifications have been performed, TANGO App Store creates an Access Token for the user, so he can use it to access services in the portal in the future.
16. The Wallet receives the Access Token and saves it temporarily to be able to request services and technology offerings from the TANGO App Store portal.
17. The TANGO App Store portal refreshes the page and displays the services and technology offerings available to the user.
18. The user is thus able to inspect a list of technology offerings from the TANGO ecosystem.
19. The user selects the desired technology offering and communicates his request with an HTTP POST request that includes also his Personal Credential Manager (PCM) URI that will be later used from services in the TANGO ecosystem in order to authenticate with the user.
20. After receiving user's selection of a TANGO technology offering, the Portal starts the communication with the TANGO Management Connector that will orchestrate the deployment of the technology offering and its associated service chain. In this step, an authentication process, similar to the one described in steps 4-16, takes place between the Portal and the Connection Manager of the TANGO Management Connector.
21. After successful authentication, the Portal sends the information for the user's PCM URI via an HTTP POST request.
22. The Connection Manager of the TANGO Management Connector acknowledges the successful reception with an HTTP 200 response.
23. Afterwards, the Portal communicates the selected technology offering to the Blueprint Repository module, so that the corresponding blueprint can be retrieved.
24. The blueprint is retrieved from the repository and is passed to the Deployment Orchestrator module for the orchestration of the deployment procedure.
25. The Deployment Orchestrator instantiates and/or configures the TANGO services that are part of the service chain described in the blueprint file. Service configuration also includes the definition of the authentication endpoints (PCM) of each other service that exists in the service chain.
26. After the instantiation of the services, each service authenticates with the other service that will need to interact with during the lifecycle of the deployment. In this step, the technology offering service authenticates with the Device Continuous Authentication (DCA) service.
27. In this step, the DCA service authenticates with the AutoML/MLOps services through their Connection Managers with the same process as described in steps 4-16.
28. Response of successful authentication is returned from AutoML/MLOps service to the DCA service.
29. Response of successful authentication is returned from DCA service to the technology offering service.
30. Response of successful instantiation of all associated services in the service chain is returned from the technology offering service to the Deployment Orchestrator.
31. The Deployment Orchestrator will need to configure the user's Connector, so an authentication process similar to the one described in steps 4-16 takes place.
32. The Deployment Orchestrator continues to execute the instantiation/configuration recipe and instructs the user's Connector through an HTTP POST request via its respective Connection Manager to download the DCA client image and the Exploratory Data Analysis Engine (EDAE) image, part of the requirements list for the DCA service, from the TANGO App Store image repository. Additionally, it passes a configuration file that will be used in

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	140 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

order to configure the DCA client and EDAE instances, enabling the discovery of endpoints of the existing TANGO services.

33. The Connection Manager of the user's Connector downloads the images from the image repository of the App Store using the Access Token provided earlier in step 16.
34. The DCA client and EDAE images are downloaded and deployed locally, on the user's connector, as they are, respectively, required for monitoring and processing user device's network and hardware metrics.
35. After the successful deployment of DCA client and EDAE, the Connection Manager notifies the Deployment Orchestrator that it is ready, and at that time the role of the Deployment Orchestrator ends.
36. In this step, as the DCA client is up and running and wants to establish a communication channel with the DCA service.
37. The Connection Manager of the user's Connector initiates an authentication process with the DCA service Connector as in steps 4-16.
38. After successful authentication, the DCA client registers itself and the user's device to the DCA service.
39. The DCA service, then, sends an HTTP POST request to the client instructing it to collect the type of data that is required, and it is available for this specific device with a specified frequency that is appropriate for this type of device.
40. The DCA client starts polling the device for the required data.
41. Raw data is returned and stored.
42. The DCA client sends the raw data to the EDAE module that is running on the same device for pre-processing, that will render the data in a pre-defined, standard format and handle invalid and undefined values.
43. The pre-processed data is returned to the DCA client.
44. As the DCA client will feed the data to the AutoML/MLOps service for classification of the device's behavior, the user's Connector must be authenticated with the AutoML/MLOps service. As such, a request to the Connector's Connection Manager is issued for initiating the authentication process.
45. The user's Connection Manager authenticates with the AutoML/MLOps service in the same manner as in steps 4-16.
46. Having acquired an Access Token, the DCA client issues a request for a specific ML service, i.e. a specific pre-trained ML model, to which it will feed the data and receive a classification prediction for characterizing device's behaviour as normal or abnormal.
47. The AutoML/MLOps service sends a request to the authorization registry (Usage Control module) of the Connector in order to retrieve the policies associated with the user's device.
48. The usage policy is returned, specifying whether the user's device has access to the service and what are the terms of usage (i.e., which model can be used, how many requests per minute, etc.).
49. The AutoML/MLOps service returns these terms to the DCA client as part of the response to the ML service request.
50. After the successful response, the DCA client initiates the data exchange of the pre-processed data with the AutoML/MLOps service by setting up a data stream using a standard data streaming software (Apache Kafka (Msg Router)).
51. The actual data exchange takes place.
52. The receiving end of the data stream in the AutoML/MLOps Connector, feeds the data to the actual ML service.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	141 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

53. The ML service/model returns a classification prediction that characterizes the device's behavior as normal or abnormal.
54. The verdict is returned to the DCA client and at that time, a data transaction is deemed completed.
55. With the completion of the data transaction, the logging of transaction's details takes place locally in both ends of the data stream, in the respective Audit Log module.
56. In this step, the logs of the transaction are communicated to the TANGO Management, Clearing House module, which may consist of one or multiple blockchain nodes or a centralized service.
57. In this example, we assume that the verdict of the ML model was of an abnormal behavior from the device. In such case, the DCA client will communicate this detection to the DCA service.
58. Sequentially, the DCA service will notify the technology offering service of the abnormal behaviour from the user's device.
59. In this step, the technology offering will stop offering its services to the device and initiate again an authentication procedure for the user's device by creating an Authentication Request.
60. The Authentication Request is returned to the user's Wallet. Along with the Authentication Request token, a redirect URI is returned that indicates the endpoint where the SIOP shall deliver the authentication result.
61. The rest of the authentication process, as described in steps 6-16, takes place.
62. If the authentication is successful, then normal service offering is resumed.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	142 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

8 TANGO Integrated Platform [T6.1-2]

8.1 Description of production environment

The TANGO integrated platform will incorporate the individual components and services described in sections 3-6 into an end-to-end system, ensuring the interoperability among the components towards covering the required functionalities of the different use cases. In principle, it will be the instantiation of the TANGO architecture described in section 7, also covering deployment aspects relative to the alignment with IDSA and GAIA-X architecture principles [45], [46].

To this end, software containerization, microservices and container orchestration technologies (e.g., Docker, Kubernetes) will be used to take advantage of the features that such technologies natively support with respect to:

- ▶ **Software isolation:** Containerization of software components allows to isolate the components from their execution environment, leading to containerized applications being totally separated from other applications running at the same hosting environment. In this context, software isolation through containerization is one of the fundamental requirements of IDSA reference architecture [46], [47] to deliver a trusted platform where Data applications are isolated and do not interfere with each other.
- ▶ **Isolation of resources within the same cluster:** On a service cluster level, there is the requirement to ensure the isolation of a cluster, which could be a dedicated IDSA connector for example, from other clusters/ connectors. In this context, Kubernetes provides some flexibility with respect to resources allocation, offering the capability to create isolated groups of resources within a single cluster through the use of separate namespaces that can build multiple virtual clusters inside the same cluster. This can be beneficial for TANGO with respect to hosting multiple Kubernetes clusters efficiently.
- ▶ **Portability:** Software containers allow developers to “ship” their applications across environments with minimal common requirements (e.g., Linux host OS). This is an important feature to be exploited for deployment within different execution environments of TANGO (e.g., instantiation of a service in a core cluster, or in peripheral (use-case) specific clusters/connectors according to dynamic configuration requirements).
- ▶ **Scalability and load balancing:** Software containerization and container orchestration technologies natively support the needs for automatic scaling and traffic load balancing across different worker nodes.
- ▶ **Automatic deployments:** Container orchestration is strongly related to performing automatic deployments following the provided configuration requirements. As explained further below, we envision automatic deployments in the TANGO production platform to be triggered through dedicated CI/CD pipelines.
- ▶ **Secure communication among deployed services:** Communication within Kubernetes clusters is secured through HTTPS and data encryption functions.
- ▶ **Fault tolerance:** Appropriate mechanisms of Kubernetes allow to recover from any errors in a transparent manner to the end user, without affecting the user experience.

As described in section 7, a hybrid deployment approach is going to be followed consisting of:

- ▶ A core Kubernetes remote cluster (hosted at Hetzner public cloud provider) that will integrate common management services (Management Connector) of the TANGO platform across the different use cases (pilots).

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	143 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- Peripheral, use-case and/or technology offering specific Kubernetes clusters that will be mapped to corresponding TANGO connectors and will be hosted either on each pilot's premises or on isolated remote cloud environments, depending on the requirements of each pilot.

As described in section 7.4.2, a deployment orchestrator will be in charge of automatically deploying each service to the appropriate hosting environment/connector according to specific configuration following the service and end-user specific requirements.

The interconnection of services running in the TANGO integrated platform with external components either running in other remote environments (e.g., NOR components for Blockchain-based Data Storage and Sharing), or Edge and/or legacy components running at the pilot sites will also be considered through the design and implementation of appropriate and secure APIs.

It is also envisioned that the production environment that is going to be used for the TANGO integrated platform will be connected with the CI/CD software stack. The CI/CD system that will be delivered and used for TANGO in the context of WP6 will incorporate DevOps best practices for daily software development and testing operations. In this context, dedicated and isolated cluster resources are going to be used for testing the individual components' functionality and the integration among different components. Automated software build, testing and deployment pipelines are going to be implemented to support the software development lifecycle (Figure 49), ensuring the fine operation of the individual components and services within the testing environment before any updates are made at the deployment (production environment).

As shown in Figure 49, the CI/CD system will be supporting both development/testing and production deployment pipelines towards totally isolated Kubernetes environments. To this end, the Kubernetes clusters dedicated to testing the technology offerings, service chains and corresponding connectors will try to mirror as much as possible the target production Kubernetes environments and desired configurations.

Moreover, in the context of validating the compatibility of the developed components (connectors) with IDSA, GAIA-X principles and reference implementations, the potential to utilize pertinent solutions for this purpose will be examined, considering their feasibility (e.g., IDS certification mechanism [48], IDS reference test suite [49], GAIA-X labeling criteria [50], GAIA-X Digital clearing house [51]).

The deployments towards the TANGO production environment are envisioned to be triggered either through dedicated CI/CD pipelines, or through the TANGO App. Store, targeting the TANGO developer/tester and end-user sides (Figure 49). For this reason, the App. store shall be able to retrieve the required blueprints and container images through the remote code and version control Github repository and the container image registry, respectively, and feed them to the deployment orchestrator. As described in section 7.4.2, the deployment orchestrator will be able to deploy, update, or delete the different services from the target connectors/Kubernetes clusters, following the configuration provided through the blueprint files.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	144 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

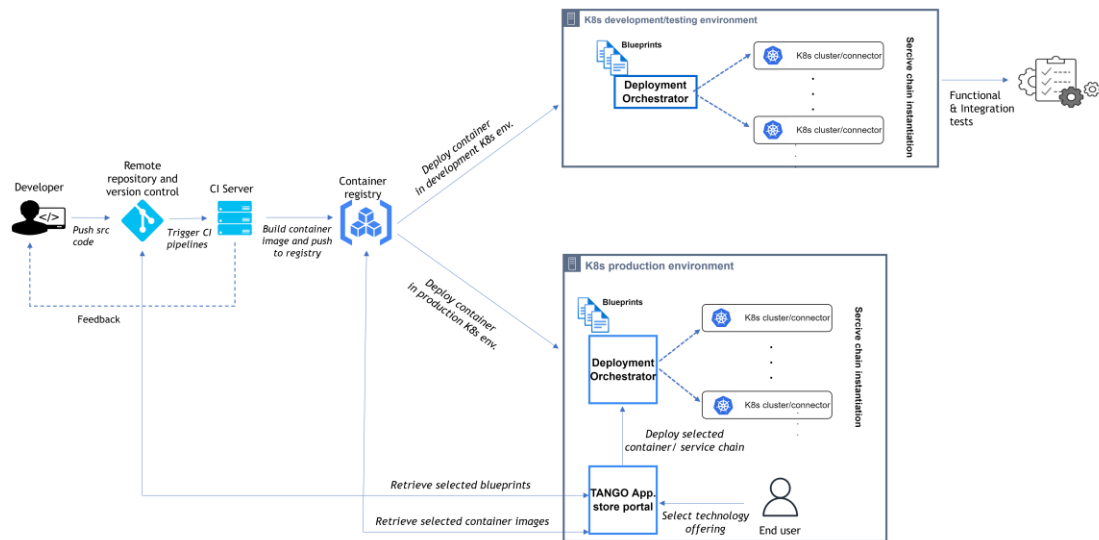


Figure 49: Automated software release pipelines and service chain deployment options

8.2 Non-functional Requirements

The non-functional requirements for the TANGO integrated platform are listed in the following table.

Table 65: TANGO Integrated platform Non-Functional Requirements (Author: INTRA, NOR)

ID	Description	Priority
NFR-COMP-TIP-001	The software components/services of the TANGO system must expose appropriate open and/or easily manageable and secure APIs (e.g., RESTful APIs over HTTPS) to ensure interoperability with other components/services.	H
NFR-MAIN-TIP-001	The TANGO system must be connected with a continuous integration and continuous delivery framework which will ensure that all the components' developments can be smoothly integrated and deployed in an operational context.	H
NFR-MAIN-TIP-002	The updates on the components of the TANGO system should not cause any sort of interruption to the operation of the platform and remain transparent to the end user.	M
NFR-PER-TIP-001	The software components of the TANGO system should be packaged and hosted in a way that will ensure efficient resource utilization.	M
NFR-PER-TIP-002	The TANGO system should natively support the capability to easily scale the deployed applications according to the dynamic requirements	M
NFR-PER-TIP-003	The TANGO system could support the building of new artifacts (e.g., container images) within seconds.	L
NFR-PER-TIP-004	The TANGO system could support the components deployment from new artifacts within seconds	L

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	145 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

NFR-PORT-TIP-001	The software components of the TANGO system must be packaged appropriately to be easily transferred and deployed to different hosting environments	H
NFR-PORT-TIP-002	The TANGO system could be able to run on any cloud or on-premise infrastructure	L
NFR-REL-TIP-001	The TANGO system must be highly available and fault-tolerant to avoid single points of failure	H
NFR-COMP-TIP-002	The TANGO system should be able to support a level of automation in decision-making with respect to the deployment environment (cluster) where an application instance should be deployed, according to specific configuration requirements.	M
NFR-SEC-TIP-001	The management interface of the TANGO integrated platform must be accessible only for admin members and protected through access restrictions and authentication/authorization measures.	H
NFR-SUST-TIP-001	Each pilot use case must be using only the absolutely necessary components of the TANGO system to operate, eliminating transactions/processes that are not mandatory	H
NFR-SUST-TIP-002	The software component executables which are integrated in the TANGO platform must be lightweight to eliminate any redundant use of hardware resources and power consumption.	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	146 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

9 Privacy, Ethical, Social and Legal Impact Assessment

9.1 Definitions

For the purposes of this section, the following definitions should be mentioned:

- I. “Personal data” means any information relating to an identified or identifiable natural person (“data subject”) 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,
- II. “Data subject” means data subject refers to any individual person who can be identified, directly or indirectly, via an identifier such as a name, an ID number, location data, or via factors specific to the person's physical, physiological, genetic, mental, economic, cultural or social identity.
- III. “Data processing” 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, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction,
- IV. “Data 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,
- V. “Data processor” means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller,
- VI. “Pseudonymization” means the processing of personal data in such a manner that the personal 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 that the personal data are not attributed to an identified or identifiable natural person,
- VII. “Anonymization” is the process of removing personal identifiers, both direct and indirect, that may lead to an individual being identified,
- VIII. “Personal data breach” means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed,
- IX. “Biometric data” means personal data resulting from specific technical processing relating to the physical, physiological or behavioral characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopy data.

9.2 Introduction

The purpose of T2.6 of the Grant Agreement “Project 101070052 - TANGO” is the conduction of a comprehensive impact assessment covering privacy, ethical, social and legal issues (PIA+) of the TANGO framework. According to the grant agreement provisions, the PIA+ is an iterative process across three phases of the project: (1) design and development of the technologies [WP2-5], (2) technical validation [WP6], and (3) evaluation of the platform [WP7]. As a whole it involves the following steps:

- 1) identify the key ethical, legal, social and privacy related themes of the TANGO platform,

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	147 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

- 2) map the data flows amongst the TANGO technologies, users, and the services with which they interface;
- 3) identify key risks associated with these data flows;
- 4) run two day-long interactive workshops and conduct interviews with relevant project stakeholders (e.g., project partners, the advisory board) and external stakeholders (e.g., technical experts, lawyers, privacy advocates, human rights experts, and proposed end-users) to validate both step 1 and 2;
- 5) engage with all partners to suggest possible technical or operational solutions, mitigation measures with special attention given to privacy, social impact, security, lawful basis, and aspects of surveillance.

In order to achieve this, a data mapping process has been implemented. Partner KEU LEUVEN has distributed to each partner a questionnaire, requesting information regarding their activities throughout the progress of the TANGO project, whether any personal data processing actions are taking place, data storage and security measures and some key ethical aspects of the process. This process is due to be repeated throughout the progress of the project in order for the data mapping process to be more accurate.

The first step of assuring compliance of the TANGO framework is the conduction of an assessment related to the development of it. As a result, apart from the assessment related to the TANGO framework, there is one been carried out beforehand related to the actions of the partners during the course project and for the purposes of it. Following the evaluation of the information provided by the partners and also considering each partners' tasks provided on the grant agreement, we have proceeded to the proper actions regarding the provisions of articles 35 and 36 GDPR [52] in order to conduct the current impact assessment. Throughout this process, a systematic description of the envisaged processing operations and the purposes of the processing, an assessment of the necessity and proportionality of the processing operations in relation to the purposes, an assessment of the risks to the rights and freedoms of data subjects and the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation have been carried out according to Article 35 par. 7 of GDPR.

Following this and during the course of the project, according to task 2.6, we are going to evaluate the architecture of the developed TANGO framework on a privacy, ethical and social related perspective. On this current section, a first attempt to proceed with this assessment has been made, taking into consideration the information we have received from the partners and the partners' respective tasks and deliverables.

We should highlight that this assessment may need to be updated throughout the progress of the project since privacy, ethical, social or legal issues may rise in the future, especially considering that, at this point of the project, many of the partners' operations have not started yet, while others still remain unaware of and cannot provide any information regarding their duties on their respective tasks and the TANGO framework has not been fully developed. As a result, it would be of outmost importance for the partners to contact those in charge of carrying out this task in order to provide them with the necessary information, whenever more information is available, since the data mapping process is very crucial in order to assure GDPR compliance and data security and integrity.

9.3 Data mapping process

Data mapping is crucial to the success of many data processes. In the scope of data protection regulations, the purpose of data mapping is firstly to consider whether personal data, as defined in the GDPR, is being processed and if so, to further examine other factors related to data protection principles such as for example the lawfulness of the processing, the respect of the

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	148 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

subjects' rights and possible threats to the subject's rights and freedoms related to any act of processing.

In order to achieve this, the TANGO Data Management Plan Questionnaire, Version 1.0, was distributed to the partners on the 30th of November 2022, asking them to provide information about the data expected to be processed during the project, how such data would be handled, and how the FAIR principles would be taken into consideration. In the questionnaire there are also questions related to the handling of research outputs other than data, as well as allocation of responses and ethics matters. As a response from our partners, we received the following questionnaires:

- eco_TANGO_D1.2_Data Management Plan Questionnaire_v1.0
- TANGO_D1.2_Data Management Plan Questionnaire_EGI_v1.0
- TANGO_D1.2_Data Management Plan Questionnaire_EXUS_v0.1
- TANGO_D1.2_Data Management Plan Questionnaire_IDSA
- TANGO_D1.2_Data Management Plan Questionnaire_SQD
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0 _ MET
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0 DBC
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0 SQUAD
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_AHOP
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_ANY SOL
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_ATOS
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_CEA
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_CESGARDEN
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_FMAKE
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_FN
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_Fraunhofer
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_IDIADA
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_INTRA
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_KUL
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_NOR
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_QBE
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_SVI
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_T3.2_Trustworthy_Data_Sharing
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_T8.5_Standardization
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_TUD
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_UMU
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_UoG
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_UoM
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_UPRC
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_VISAR
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0_VTT
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0-ABILAB
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0-FUJITSU_LU
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0-LST
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0-UTH
- TANGO_D1.2_Data Management Plan Questionnaire_v1.0-XLAB
- TANGO_D1.2_Data Management Plan Questionnaire_NTT DATA Italia v1.0

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	149 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

After the evaluation of the answers received, we were able to witness whether any personal data is being/will be processed and if this is the case, whether further measures should be taken regarding this. The answers are also an indicator of the knowledge that TANGO partners have related to data protection principles, the level of their compliance in relation to their project activities and the measures they have implemented (or plan to implement) in order to achieve data security and integrity. The answers received in response to the questionnaire constitute initial input and data management processes, including obtaining further information about partners' data-related practices, will iteratively continue during the project.

Moreover, by resorting to the grant agreement guidelines regarding the work packages, we were able to acquire information related to the basic operations of each partner. Following this procedure, we contacted each of the partners whenever necessary for additional information, while also organizing interactive workshops and conducting interviews with relevant stakeholders.

Furthermore, the feedback received for the purposes of conduction of this deliverable was very crucial in order to identify the implementation of the proper ethical and privacy related legislation during the development of the TANGO framework. Factors such as the “privacy by design” and the use of Artificial Intelligence technology (AI) were key and were thoroughly examined to this point.

Although some partners could not provide any information regarding their duties, we have reached a conclusion regarding the personal data that may be collected, generated, re-used and/or in general processed during the progress of the project. During the later stages of the project, it is possible that we could identify more data processing activities given that for now, as stated above, many of the partners do not have the necessary information regarding their duties and are not in a position to provide us with any information regarding further data processing acts that will take place in the future.

9.4 Description of the data processing activities during the development of the TANGO framework

Following the data mapping process, we were able to reach some conclusions regarding the currently running data processing activities until the publication of the current deliverable. On this subsection, the currently traceable processing activities will be presented, while referring to the nature of this personal data and the way they have been collected. Moreover, since the pilots are due to start after the publishing of this deliverable, we can predict some of the data processing activities that are due to take place in the future. As a result, as mentioned above, we should highlight that the DPIA+ is a repeated process that requires constant communication between the partner in charge of this task and the rest of the partners.

Following a careful and thorough review of the answers provided by the partners, we can infer that processing of personal data takes or will likely take place in the cases mentioned below. We also present some information regarding the data processing activities related mostly to the nature of the personal data, while presenting some key GDPR provisions. A thorough analysis regarding the lawfulness of the data processing will be presented on the next section. It is noteworthy that the list is not restrictive, rather subject to changes and updates as the project proceeds since we are in its early stages and the piloting activities have not started yet.

On the table below, the personal data processed by every partner are presented, with reference to the cases in which special categories of data are included, the necessity of those processing operations in regard to the project purposes and finally the processing period. In some cases, predicted data protection activities are referred also, especially in terms of the processing activities that will take place during the pilots.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	150 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 66: Personal data processed per TANGO partner

Partner	Personal data processed	Special categories of data	Purpose of processing	Processing period
ECO	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
EGI	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country, ip address) b) Personal and contact information of stakeholders (name, surname, email address, telephone number, place of residence/work, country, ip address) c) Behavioral data of the stakeholders (usage data such as websites, services, social media and log in timestamps)	-	a) Cooperation for the purposes of the project b) May be needed during the process of creation of the TANGO platform c) Will be used according to the purposes of the project in order for the platform to provide the stakeholders with the intended service according to the grant agreement	a) Documents generated will be preserved for at least 5 years. b) Personal and contact information for at least 1 year in case there is no request to delete them after the end of the project c) The rest until the end of the project
EXUS	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) b) Personal and contact information of the users of the participants on the pilots regarding the online store and the autonomous vehicle usage such as name, surname, address, credit card details, location, device ID, cookies) c) Behavioral data of the participants (transactional data, user interaction with various store functionalities, device information, likes on social channels, reviews)	-	a) Cooperation for the purposes of the project b) Performance of data analytics according to the purposes of the project in order for the platform to provide the stakeholders with the intended service according to the grant agreement	Until the end of the project
IDSA	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
SQD	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	151 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

MET	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
DBC	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) b) Personal and contact information of the participants on the project (name, surname, email address, telephone number, place of residence, country)	-	a) Cooperation for the purposes of the project b) Creation and the presentations of the deliverables, organization of workshops	Until the end of the project
SQUAD	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
LIC	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) b) Personal and contact information of the participants on the project (first name, last name, e-mail, place of residence/work, role in the organization)	-	a) Cooperation for the purposes of the project b) Organization of the communication and dissemination activities conducted for the project	Until 90 days after the end of the project
AHOP	-	-	-	-
ANYS	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) b) Participants' on pilot 7.2 preferences (sex, age, nationality, gastronomic preferences, leisure preferences)	-	a) Cooperation for the purposes of the project b) Collection, utilization and manipulation of the data in order for them to be used properly during the progress of the 7.2 pilot and thus help with the creation of the platform	Until the end of the project
ATOS	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) Personal and contact information of the participants in the project (id, name, surname, date of birth – age and personal identifier)	-	Cooperation for the purposes of the project	Until the end of the project
CEA	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
CESGA	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	a) Cooperation for the purposes of the project	Until the end of the project

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	152 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

	b) Participants' on pilot 7.2 preferences (sex, age, nationality, gastronomic preferences, leisure preferences)		b) Collection, utilization and manipulation of the data in order for them to be used properly during the progress of the 7.2 pilot and thus help with the creation of the platform	
FMAKE	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
FN	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
FHG	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
IDIADA	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) b) User identifier	-	a) Cooperation for the purposes of the project b) As a result of the connection of the platform to the vehicles	Until the end of the 7.4 pilot
INTRA	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
KUL	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
NOR	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) b) Personal and contact information of the participants on the pilots (name, surname, middle name, birth name, residence address, country, email address, telephone number, place of work- and work-related information, salary and tax related personal data, data regarding their bank accounts and their transactions with the use of the banking sector)	National ID, passport and/or VISA information and citizenship information of the participants	a) Cooperation for the purposes of the project b) Data collection in order for them to be used for the tests of the platform during task 7.5 (banking) and according to the purposes of the project	Until the end of the project
QBE	a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country) b) Behavioral data of the participants in terms of a person's movement, device usage and transactions	Facial biometric data of the participants	a) Cooperation for the purposes of the project b) User authentication methods for the purposes of the task 7.5 (banking)	Until three years after the collection of the data in case there is no request for erasure. If there is such a request, the data will be deleted instantly

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	153 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

SVI	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
DUT	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
UMU	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
UOG	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
UOM	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
UPRC	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
VISA	<p>a) Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)</p> <p>b) Personal information of the participants (name, surname, employment and educational history, details of any previous visas or residence permits)</p>	<p>Nationality and country of origin of the participants, family information (names and dates of birth of the family members),</p> <p>information about any criminal record or security concerns and medical and health information, regarding any disabilities or special needs</p>	<p>a) Cooperation for the purposes of the project</p> <p>b) For the purposes of creating and achieving the goal of the platform related to the task 7.6 (public organizations – visa applications)</p>	<p>a) Until the end of the project</p> <p>b) Regarding the visa application related information, the data will be stored for as long as required by the German law</p>
VTT	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	154 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

ABI	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
LSTECH	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
UTH	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
XLAB	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project
NTT DATA Italia	Personal and contact information of the project partners (name, surname, email address, telephone number, place of work, country)	-	Cooperation for the purposes of the project	Until the end of the project

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	155 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
		Status:	Final		

Currently, since we are in the early stages of the project, there are several partners unaware of the possibility of carrying out any personal data processing operations since their duties are mostly related to the conduction of the pilots. As a result, we remain in constant communication with all the partners for the purposes of noting all the data processing operations, updating the current assessment and assuring proper monitoring of them.

Moreover, due to the technological aspect of the project and the fact that many of the partners are technological partners with their task being the development of software, it is predicted that some of them will not process any personal data of participants to the project and only process personal and contract information of the project partners that they cooperate closely for the purposes of communication and coordination related to the project.

9.5 Lawfulness of the data processing

On this section, the lawfulness of the data processing activities is presented. For starters, a reference to the basic GDPR provisions regarding data processing has been made, presenting the basic principles relating to processing of personal data and the rights of the data subjects.

Following this, the implementation of those provisions has been presented, with reference to the measures adopted by the partners. It should be highlighted that the lawfulness of the data processing is the most crucial part of data processing since it justifies any data processing activities, thus assuring data security and integrity.

The general rule in data processing is that any data processing operation is deemed unlawful unless the conditions laid down on the GDPR provisions are being met.

9.5.1 Key provisions of GDPR

A necessary condition in principle to achieve the lawfulness of the processing of a data subject's personal data is the cumulative fulfilment of the following the fundamental principles relating to the processing of personal data, according to article 5 GDPR. These are mandatory provisions which are necessary to be applied cumulatively and compliance with them must be examined at each individual stage or the processing operations. Throughout the progress of the TANGO project, the partners must assure that the personal data is:

- i. processed lawfully, fairly and in a transparent manner in relation to the data subject (“lawfulness, fairness and transparency”),
- ii. 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”),
- iii. adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed (“data minimization”),
- iv. 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”),
- v. 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 organizational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject (“storage limitation”)

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	156 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

- vi. processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures (“integrity and confidentiality”).

The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 (“accountability”).¹

Furthermore, according to article 6 of GDPR, processing shall be lawful only if and to the extent that at least one of the following applies:

- i. the data subject has given consent to the processing of his or her personal data for one or more specific purposes,
- ii. 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 contract,
- iii. processing is necessary for compliance with a legal obligation to which the controller is subject,
- iv. processing is necessary in order to protect the vital interests of the data subject or of another natural person,
- v. 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 controller,
- vi. 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 child².

Finally, according to article 9 of GDPR, the TANGO partners acknowledge that processing of special categories of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation is only allowed in case:

- i. the data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject,
- ii. processing is necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject in the field of employment and social security and social protection law in so far as it is authorized by Union or Member State law or a collective agreement pursuant to Member State law providing for appropriate safeguards for the fundamental rights and the interests of the data subject,
- iii. processing is necessary to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent
- iv. processing is carried out in the course of its legitimate activities with appropriate safeguards by a foundation, association or any other not-for-profit body with a political, philosophical, religious or trade union aim and on condition that the processing relates solely to the members or to former members of the body or to persons who have regular contact with it in connection with its purposes and that the personal data are not disclosed outside that body without the consent of the data subjects,
- v. processing relates to personal data which are manifestly made public by the data subject,
- vi. processing is necessary for the establishment, exercise or defense of legal claims or whenever courts are acting in their judicial capacity,
- vii. processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data

¹ Article 5 GDPR

² Article 6 GDPR

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	157 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject,

- viii. processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3,
- ix. processing is necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy,
- x. processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.³

The TANGO partners are obligated in case it is identified that the processing operations do not comply with these basic principles - requirements of the GDPR, as detailed above, to refrain from any processing activities as it is deemed unlawful pursuant to the general rule of data processing. They are also aware that any processing operations of personal data provided to them during the progress of the project for purposes other than the ones stated in the grant agreement is forbidden and that, even if any other legal basis justifies the forementioned processing operations, they are obligated to inform the data subjects for their operations according to arts. 12-14 GDPR.

9.5.2 Implementation of those provisions to the data processing for TANGO

In application of the principle of accountability, the data controller shall bear the burden of proof for any action taken by it which is related to the lawfulness of the data processing and becomes accountable for his actions before the competent supervisory authorities and the data subjects. As a result, the necessary organizational and operational measures have been implemented during the data processing operations for the purposes of the TANGO project, thus assuring lawfulness of the data processing activities and consequently, data security and integrity.

According to the feedback received, all the project partners who have engaged or are due to engage in any processing operations of personal data, have made sure that the provisions of art. 6 GDPR are being met.

For the purposes of the project in terms of cooperation between the partners, each one of them has access to specific personal and contact information of the rest of the partners - data subjects. The forementioned information is presented on Table 66 and each partner has access to them through the contact information provided on the files “TANGO_contact list_v9.1.xlsx”, for which each partner has to enter a specific code in order to have access to the content of the file, and “TANGO_contact list_17042023.pdf”, located in the OwnCloud server. Partners can also have access to such personal data of project partners through the deliverables submitted and the various events or workshops that take place throughout the progress of the project. The legal basis for such data processing operations is “performance of a contract” (grant agreement) according to art. 5 par. 1b’ GDPR and they are limited only to the necessary ones related to the duties of each partner on their respective task – work package.

³ Article 9 GDPR

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	158 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Furthermore, for the purposes of the development of the TANGO platform on the various sectors provided on work package 7, it is expected that personal data of the participants to those pilots will be processed, according to each task – socioeconomical sector.

On the pilot related to the visa applications procedure for instance, it is expected that the participant has to provide the partner with all the necessary personal data that are being provided by the data subject in such cases, in order for the platform's function to be successfully tested on an environment similar to the usual procedures that are followed in such cases. It is though noted that the partner in charge of this task – pilot is an entity that either way proceeds to such processing operations according to their usual operations and the personal data processed are the ones that are appropriate and need to be collected and in general processed while carrying out their duties. The forementioned personal data are collected and processed in general after the data subjects provide their consent, according to art. 6 par. 1a GDPR, remain confidential and only the pilot – task leaders have access to them.

Regarding the consent provided by the data subjects, it is noted that the partners follow all the provisions related to this legal basis of processing. The partners assure that the consent is freely provided, specific to the processing operations and an unambiguous indication of the data subject's wishes. During this process, the data subjects receive the necessary information regarding the data controller (brand name of the data controller, headquarters location, contact information, information related to the DPO or any contact person for privacy related matters), the personal data that will be processed, the processing operations and the purpose of such processing operations, with a reference to the TANGO project and the specific pilot(s) for which the data will be processed. The data subjects are also thoroughly informed about their rights and the process of withdrawal of consent and receive any other information they wish regarding the processing operations.

Moreover, another pilot in which personal data will be processed is the one related to banking. Various partners, such as QBE and NOR, will proceed to processing operations of personal data in order for the platform to be developed and properly function. For the purposes of user authentication and security, QBE will require from the users to provide facial biometrical data and have access to behavioral data (for example location, device usage and transactions), while NOR will also collect and in general process personal data related to those usually provided by the data subjects to a bank. The data subjects will provide access to the forementioned personal data to the partners following the provision of the necessary consent, according to art. 6 par. 1a GDPR. The data subjects will be able to refrain from the provided consent at any time and upon request, their personal data will be instantly erased.

Apart from the forementioned measures implemented to assure that the partners proceed to the data processing activities lawfully and according to article 6 GDPR, there are further measures that have been implemented that the processing activities are restricted to the absolutely important ones.

The partners as data controllers or processors have made sure that only the authorized partners have access to the personal data of other partners or the participants in the project. For starters, the access of the partners to personal information of the rest of the partners is limited with the implementation of a code needed to be filled in for the excel file mentioned above to be opened and processed. Moreover, access to the personal data of the participants on the pilots will only have the partners participating to each pilot while some deliverables including personal data and deemed sensitive in the grant agreement, thus blocking any unauthorized access.

9.6 Implemented measures related to the data subjects

On this subsection, the respect to the rights of the data subjects related to GDPR is stated. The key provisions regarding the rights of the data subjects are being mentioned in detail, with a detailed reference to the measures being implemented in order to assure that the data subjects remain in control of their personal data and are able to exercise their rights effortless and without any undue delay.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	159 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

9.6.1 Key provisions of GDPR

The respect for the rights of the data subjects is of paramount importance. Those are presented on Articles 12-22 of the GDPR and refer to the:

- i. **“Right to be informed”** regarding anything that is related to the processing of personal data (arts. 12-14 GDPR),
- ii. **“Right of access”** to the personal data and any other information related to the data processing activities (art. 15 GDPR),
- iii. **“Right to rectification”** of inaccurate personal data or completion of incomplete personal data (art. 16 GDPR),
- iv. **“Right to erasure” (“right to be forgotten”)**, according to which the data subject could achieve the deletion of his/her personal data under certain conditions (article 17 GDPR),
- v. **“Right to restriction of processing”** of the personal data under certain conditions (art.18 GDPR),
- vi. **“Right to data portability”**, according to which the data subject can receive the personal data concerning him/her, which had been provided to a controller, in a structured, commonly used and machine-readable format and transmit them to another controller without hindrance from the controller to which the personal data had been previously provided (art. 20 GDPR),
- vii. **“Right to object”** at any time to processing of personal data concerning him/her (art. 21 GDPR),
- viii. **“Right to object to the automated individual decision-making, including profiling”** (art. 22 GDPR)⁴.

Article 12 of the GDPR sets out the procedure necessary to be followed when exercising the rights of the subjects. The data subject shall submit a request to the controller who shall, in accordance with the third paragraph of Article 12 of the GDPR, reply to the request without delay and in any event within one month. That period may be extended by a further two months if necessary, taking into account the complexity of the request and the number of requests. In the latter case, however, the controller shall inform the data subject of that extension within one month of receipt of the request and of the reasons for the delay. Where the data subject makes the request by electronic means, the information shall be provided, if possible, by electronic means, unless the data subject requests otherwise.

Furthermore, in article 23 of the Regulation some restrictions on the exercise of the rights of the subjects are presented. However, it is necessary to stress that these restrictions apply only in cases where they are applicable in accordance with the GDPR and the core of the rights is not infringed but the protection of a higher legal interest is sought in the context of a democratic society.

9.6.2 Currently implemented measures

The partners of the TANGO project have made sure to implement the necessary measures in order to assure that the data subjects remain in control of their personal data and can exercise their rights effortlessly and without any undue delay, thus eliminating any risks affecting the data subjects’ rights and freedoms.

Firstly, in case any data subject wishes to exercise his/her rights, each partner has made sure to appoint a contact person for the data subject to communicate with. Upon receiving any request from the data subjects, the contact person reacts instantly and according to the forementioned GDPR provisions, providing the data subject with the necessary information in order to fulfill the request and if needed, he contacts the authorized legal partner of the project for further information regarding this matter. The authorized legal partner of the project is DBC and a partner in need of assistance can contact DBC through the e-mail address legalint@diadikasias.gr.

⁴ Articles 12-22 GDPR

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	160 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Moreover, through the provisions related to the Horizon Europe projects but also the data management plan issued for the TANGO project, the basic framework has been stated regarding the rights of the data subjects. The partners are properly informed regarding this matter and the way they should react in case of an application from the data subjects. They have the necessary background to respect those provisions and react properly and according to the GDPR and any other related provisions, thus assuring that data subjects' rights and freedoms are not at risk.

9.6.3 Actions in case of a data breach event

A data breach event, whether due to malicious action by a third party or due to negligence of any kind within an entity, is the most serious event provided for in the GDPR as it entails a circumvention of the principles of it in case of negligence and an infringement of data protection and the rights of data subjects. In articles 33 and 34 of the GDPR, the necessary actions to be taken by the controller before the personal data protection authority and the data subjects in order to limit losses as far as possible are provided.

In compliance to the GDPR provisions regarding this matter, the partners of the TANGO project are appropriately qualified to proceed to the necessary actions in order to minimize or eliminate the consequences of such an event. Each partner accepts his/her obligation to notify within 72 hours the authorized, in accordance with Article 55, supervisory authority and the data subjects, according to articles 33 and 34 GDPR, documenting in detail the personal data breach and the facts relating to it, the personal data involved, the consequences of the event and the corrective measures implemented.

9.7 Key risks

On this subsection, the potential risks and issues that may occur on data processing activities regarding both data security and integrity and the data subjects' rights and freedoms will be presented. The use of new technologies is always a challenging factor on cases like this and in order to be more accurate with this part of the assessment, we have also contacted the partners in order to consult them regarding the possible risks that they may face during their data processing operations.

As it is foreseen on the GDPR, the DPIA is the safest way to predict such events in advance in order for the data controllers to be preemptively prepared to anticipate them. As a result, the help of the partners is much appreciated since their expertise has helped us identify any possible data protection related issues they may face in the future. Of course, as stated above, this assessment will be subject of updates since privacy, ethical, social or legal issues may arise in the future.

Some events that could be predicted at this stage and could arise data protection related risks – issues, are the following:

- i. Unauthorized access to the personal data as a result of
 - a) Data leakage through the installation of a third-party software
 - b) Unauthorized physical access to the servers of an entity
 - c) Unauthorized disclosure
 - d) A fraudulent event
- ii. Malicious attacks to the TANGO server and/or the hardware and software that the partners use for their tasks with the use of mechanisms such as Malware, Phishing, Ransomware, Denial of Service, Man in the Middle or any other related ones.
- iii. Restriction of access to a hardware device (computer, mobile phone etc.) or software because of a malfunction, theft or any other instance
- iv. Accidental data loss or destruction
 - v. Use of software or hardware with insufficient security architecture
- vi. Improper erasure of the personal data

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	161 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

- vii. Use of insufficient means of security
- viii. Storing of incomplete or inaccurate personal data
- ix. Human error
- x. Loss of data because of a natural disaster

The forementioned events are commonly addressed ones and as a result the TANGO partners have proceeded to all the necessary actions and implemented the necessary technical and organizational measures, as stated on the next section, according to the GDPR, the guidelines of the European Data Protection Board and the independent supervisory authorities in order to assure data security and integrity.

9.8 Technical and operational security measures

Ensuring the security and integrity of personal data, both at the stage of processing and at the stage of subsequent storage, is one of the most critical obligations of the data controller. In the context of the climate of protection established by the GDPR that obligation is emphasized in various parts of the Regulation, with the latter even proposing some protective measures aimed at assuring data security.

9.8.1 Key provisions of GDPR

The GDPR has adopted a “risk-based approach” for each case. According to this, each case should be evaluated according to the degree of risk to the rights of the data subject and the corresponding imposition of obligations. In other words, the necessary technical and operational measures that each data controller implements should be on the level required in order to anticipate the potential risks that he/she may face while proceeding to the data processing operations.

Moreover, according to article 25 of GDPR, the data controller should, *taking into account the 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 natural persons posed by the processing, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organizational measures, which are designed to implement data protection principles (“security by-design”) and implement appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed (“security by-default”).*

9.8.2 Currently implemented security measures

Taking those into consideration, our partners have implemented the following technical and operational measures provided on this subsection. Those are considered to be appropriate and capable of anticipating any predictable risks and issues, according to their technical expertise and the related security guidelines, thus assuring that the current assessment regarding the forementioned data processing operations is efficient, comprehensive and in compliance to the GDPR.

9.8.2.1 Organizational measures

On this subsection, the organizational measures implemented by the partners are presented. The partners have implemented specific organizational measures to assure data security and integrity during the operation of the data processing activities, such as restriction of access protocols, proper training, data management plans and various procedures regarding their responsibilities according to the GDPR.

The need of proper training and spread of awareness regarding data protection matters is highlighted in many instances by both the European Data Protection Board and the independent supervisory authorities of the members – states. The participant representatives of the partners possess the necessary knowledge and awareness regarding data protection related issues, factors that had been taken into consideration

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	162 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

during each partner's decision process regarding their representatives - participants to the project. Each one of them on their respective tasks, taking also into consideration their background (mostly technical, related to informatics), prove their respect to the GDPR provisions through the development of software and/or hardware following a "personal data protection" architecture.

An example of this training is the procedure that partner INTRA follows in order to spread awareness related to privacy related matters. The organization Netcompany-Intrasoft has launched the "Information Security Awareness e-learning", a unique in-house initiative designed to empower employees with essential skills and knowledge in Information Security. Through this program, Netcompany-Intrasoft aims to further reinforce the company-wide culture that values and protects valuable information, ultimately enhancing customers' trust to INTRA's services and products.

Moreover, the partner DBC has proceeded to many actions related to the spread of such awareness. Through interactive workshops, the related deliverables, communication with the partners, the interaction through the project meetings and their interventions in any case any relevant issue arises, they have provided the partners with the necessary feedback and guidelines to be followed, thus assuring the spread of awareness regarding data protection matters.

Apart from this, the implementation of "access to personal data policies" has been a priority for all the partners from the start of the project. Each partner has for starters appointed the people participating to the project, limiting any exposure to personal data exclusively to them. As a result, other colleagues of them, who either do not participate to the project or are in charge of carrying out other tasks of it, do not have unauthorized access to personal data. Their colleagues who participate on other tasks of the project have access only to the personal data they need in order to carry out their duties, according to the purposes of the project, the grant agreement and the data management plan. In general, access to specific personal data (and data in general) is provided according to the purposes of the project and the task that each of the participating body has to carry out, while simultaneously evaluating the relevance of the personal data to the forementioned tasks to be carried out by the partner.

Furthermore, the partners have adopted rules regarding the management of the data in general. Each partner has drafted policies with which the data are separated to common data and personal data according to GDPR, while the latter are also divided to a) those related to special data according to article 9 GDPR and b) the rest of the personal data.

This action is the first step that ensures that a stricter policy is implemented regarding personal data. As a result, the partners adopt rules related to the GDPR provisions in order to assure data security and integrity. Whenever needed, the personal data is updated through checks that take place on a regular basis, policies that restrict access to the personal data only to specific people are adopted and the access and in general processing of personal data is recorded through reports and the records of processing activities. Data transfers outside the premises of the organization – partner are allowed only if it is considered absolutely necessary for the purposes of the processing activities, after a permission is provided and are recorded. Rules have also been implemented regarding the archives and the access to them, with restriction to access policies and monitoring.

Regarding the erasure process of the personal data, the partners have also set out specific rules and procedures. By the time the personal data are no longer needed, a legitimate request by a data subject to delete them has been made or there is no other legal binding obligation to continue the processing operations, such as storage for example, the partners have implemented the necessary organizational and technical measures to assure that the personal data is completely deleted by the partner's infrastructure. Since, as mentioned above, the processing operations are monitored, the partner is aware of the existence of possible copies (backups for example) that may exist and proceeds to the deletion of them, assuring that the personal data is completely vanished from any software or hardware that they possess.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	163 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

The basic procedure regarding the deletion of documents that contain personal data, is the following:

a) Regarding hard copies:

- i. Compilation of the documents to be deleted on a specific folder/place in which the data controller or authorized personnel from the data controller have access. This process assures that only the planned personal data will be erased.
- ii. Destruction of such documents with the use of related hardware, such as paper shredders, located in the premises of the data controller.
- iii. Document shredding may be followed by pulping and/or recycling of the documents, only after it is assured that the personal data contained in the remains of the document cannot be identified.

b) Regarding electronic archives

For the safe destruction of data in electronic form, simply deleting the data is not considered sufficient, as this deletes only the reference to the data, while the data itself may be recoverable using special software programs. Although this action also takes place and the participants in the project make sure that there are not any files remaining in the “recycling bins”, any email attachments or in general any place in a software/hardware, further measures are implemented to ensure safe deletion of the documents containing personal data.

As a result, the destruction of data stored on rewritable media (e.g. hard disks, floppy disks etc.) is taking place by the implementation of alteration technics, for example by overwriting them with random characters. The alteration can be done by using special programs (file erasers, file shredders, file pulveritizers). In the case of everyday data destruction, an alternative way of destruction taken into consideration is the formatting of the substrate material (format). Finally, in the case of scheduled destruction of all personal data, the possibility of physical destruction of the substrate material itself is also examined (e.g., by crushing, pulverization or incineration, subject to specific provisions on special waste management / environmental protection) as an alternative means of destruction.

The final step of the personal data erasure process is its recording in the processing records and/or the conduction of a “protocol of erasure”.

Finally, regarding any personal data breach event, a specific procedure has been conducted following any such event. For starters, the relevant framework has been presented to the partners on the data management plan deliverable, thus providing the partners with the necessary information regarding their obligations. Also, as a result of those provisions, each partner follows a specific plan, consisting at least of the following steps:

- i. Realization of the data breach event and assertion of the severity of it. The data controller figures out whether any personal data is involved and whether the breach is still occurring. If the breach is still occurring, the data controller should implement the necessary measures according to the nature of the event in order to stop further data loss.
- ii. Immediate action to mitigate the consequences of the data breach event, by implementing the necessary measures to recover any data loss where possible and limit damage caused.
- iii. Notification without undue delay and, where feasible, not later than 72 hours, of the competent supervisory authority.
- iv. Notification of the data subjects in case of data breach events likely to result in a high risk to their rights and freedoms.
- v. Notification without undue delay of their legal department and/or the authorized legal partner of the project (DBC) for further guidelines.
- vi. The decision taken as to the reasons why a data breach is either reported or not reported is documented by the data controller.
- vii. A review of both the causes of the breach and the effectiveness of the response is conducted in order to witness any security related issues that may have arisen as a result of the data breach.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	164 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

The partners constantly proceed to checks and if needed updates of the contact lists, thus assuring that their communication takes place with the right people – participants of the project. They have made sure that any partners who no longer participate to the project (in the case of TANGO, the partner participating on the banking pilot has been changed), are erased from the contact list and any TANGO related communication with them has stopped.

9.8.2.2 Technical measures

On this subsection, the technical measures implemented by the partners are also presented. The partners have adopted specific rules to assure data security and integrity during the use of their infrastructure, such as restriction of access mechanisms, back-ups and relevant software architecture.

Firstly, several restriction of access mechanisms have been implemented on the software and hardware utilized by the partners. Access to the infrastructure that the participants of the partners use for the purposes of the project is limited to themselves, as part of the organizational measure presented on the previous subsection, thus assuring that there will not be any unauthorized access to any personal data.

This is ensured with the implementation of measures such as the use of strong passwords in order for the infrastructure to be used and the user to be identified safely. The partners implement a “password policy”, in which the length, characteristics and the frequency in which the password should be changed are foreseen, while also rules regarding the prohibition of disclosure of them and the ways those should be noted are in place. The participants are obligated to use passwords with at least one upper case letter, one lower case letter, one number and one special character, while mandatory periodic password change are foreseen. They are also obligated not to reveal their password to any third parties and only do it in case it is absolutely necessary for the purposes of carrying out the partner’s duties regarding the project, following a permission from the data controller. In such case, the participant shall change the password in order to retrieve the “password’s confidentiality”.

A further mechanism is the monitoring of access. During the implementation of the software / hardware and in general means used for the purposes of the project, the partners have made sure that a basic “data protection architecture” has been implemented. Such characteristics relatable to the forementioned architecture are:

- i. the restriction of access to a user who has multiple unsuccessful access requests
- ii. mechanisms that make sure that the hardware or software does not continue operating whenever left idle
- iii. complete monitoring of the operations of hardware / software that contain information regarding the personal data with the use of log files (user identification, timestamp, system used and whether the access had been completed)
- iv. mechanisms that shall make any intervention to the software / hardware impossible.
- v. monitoring of use of the any portable devices (USBs, memory disks etc.)
- vi. the development of the necessary tools to detect any possible privacy and security issues

Moreover, the installation of security-related software has been of outmost importance for the partners. Apart from the training related measures in regard to the anticipation of cases in which a possible threat that a malicious attack may take place, the partners have also obtained and utilize state of the art security related software, such as for example ESET Endpoint Antivirus, Windows BitLocker for hard-drive encryption and VPN, in order to detect any possible threats (for example viruses, malware, phishing, ransomware, denial of service, man in the middle or any other related threats) and anticipate them, following the procedure relatable to any of these cases, which consists of the following step:

- i. Realization of the nature of the threat

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	165 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

- ii. Instant implementation of the necessary means to anticipate it. The data controller shall make sure such events will stop, any malicious software is deleted from the software / hardware and any event of such will not be repeated in the future, with the implementation of the proper security measures
- iii. Implementation of the data breach management plan if applicable

The creation of back-up files is also another measure applied. The partners specify the personal data that need to be included in back-up files according to their importance and the frequency of the processing operations. Following this specification, they determine the frequency of creation and update of those files and ensure that those back-up files remain accessible to the authorized personnel and securely stored.

A common strategy followed regarding the back-up files is the “3-2-1 back-up system”. According to this, one primary backup and two secondary copies of the data are created which should be stored in two different types of media and one of them is kept at least in a place different than the premises of the partner. Partner INTRA will be storing the back-up files on the Hetzner cloud that will be purchased for the purposes of the project.

The use of proposed security methods such as anonymization, pseudonymization or encryption of the data has also been prioritized. For starters, whenever it is possible, the personal data are anonymized in order to comply with the GDPR provisions. Most of the technical partners involved in the software development in case it is needed for their tasks to process any personal data, proceed to the implementation of anonymization technics in order to ensure that all the identifiers are erased since they do not need any, according to the GDPR, personal data for their tasks.

In case the anonymization of the personal data is not possible for the purposes of the project (when for example anonymized data will not be useful for the purposes of a specific task), the implementation of the methods of pseudonymization or encryption has been examined.

It should be noted that the contact information of the partners and the data subjects participating on the pilots cannot be in most of the cases pseudonymized since the partners having access to them (according to the specific purposes of the project) will be unable to use them for the purposes of the project.

9.8.2.3 Means of communication and data sharing

The partners have implemented the necessary measures in order to assure security of communication between them, thus ensuring personal data security and integrity. The use of the following means of communication between them was thoroughly examined during the decision process and security related factors, such as the technical and organizational measures that the platforms implement and their architecture, had been taken into account before the final decision.

The partners have decided to use the OwnCloud cloud-storage platform, an open-source software, for the purposes of storing and sharing of the project related documents, which of course may contain personal data. This platform is considered to be secure, having adopted a comprehensive and strong security architecture, ownCloud ensures the data is stored and processed securely and in compliance with the GDPR. It encrypts files, blocks illicit uploads through antivirus and illicit file access through the file firewall. In November 2022, ownCloud published version 2.0.0 "Infinite Scale", which was developed with the collaboration of the European Organization for Nuclear Research, which is also a sign of the importance that the provider shows on security matters.

9.8.2.4 Project activities supervision

The operations of each partner in regard to the project is subject of periodical supervision both internally and by the authorized legal partner of the project. This process helps the partners be aware of all the privacy related issues they may face and be aware of their responsibilities as data processors / controllers.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	166 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

DBC, as the authorized legal partners has proceeded to various actions in order to successfully map all the data processing operations and provide the partners with the necessary feedback whenever needed. For the purposes of the data management plan, the partner KU Leuven had distributed the partners with a questionnaire related to their data processing operations. The feedback received was evaluated for the purpose of the current assessment and whenever needed, further information to be provided was asked from the partners via email or any other means of communication possible.

Communication with the partners on a periodic level and especially whenever it may be important as a result of the progress of the project (for example, by the time the pilots have been started) is particularly important in order to successfully proceed to the personal data mapping process and DBC will engage to workshops too in order to ensure that.

Finally, the need for cooperation was highlighted during the meetings and in general communication with the partners. For the purposes of compliance to the GDPR, ethics and any legal provisions, the partners acknowledge their duty to communicate with the authorized legal partner of the project in order to receive the necessary guidelines and provide them with any new information related to their project tasks. Since we are currently on the early stages of the project, many partners may not have a clear view regarding their overall duties, thus being unable to provide the necessary information related to any data processing operations.

9.9 Assessment related to the architecture of the TANGO framework

Following the assessment related to the relative to the project purposes processing operations, a comprehensive one related to the already developed technology should be conducted. On this section, an evaluation of the architecture of the TANGO framework and the already developed or planned to be developed technology in terms of privacy has been conducted.

On the previous sections of this deliverable, our partners present the proposed software elements that will be developed and used on the TANGO framework and the achievement of adequate levels of security and in general personal data privacy is highlighted.

The tasks described in WP3 of the grant agreement refer to the data storage and in general sharing during the use of the TANGO framework. T3.1 refers to the use of the Fides platform, a blockchain based data/attestation sharing platform. The use of Fides platform, developed by the partner NORBLOC, is based on user's consent and provides him/her with the possibility to manage the provided consent at any time freely and without any unnecessary additional actions. It also contains access control mechanisms through verification methods upon entry and the retain of audit logs, while the use of the DnAMS (Decentralized norbloc Access Mgmt System), which is a highly elaborate key management mechanism, employing 3 layers of encryption as well as a split-seed design, and prevents data access even in cases of malicious attacks to the platform.

Moreover, the features developed for the purposes of T4.1, T4.2 and 4.3 also strengthen the restriction of access protocols. T4.1 refer to the Self-Sovereign Identity (SSI) Solution developed which will provide the user with the possibility to manage their connection credentials for the purposes of proceeding to the transactions supported by TANGO, T4.2 to the DUOS platform for physical and logical access control for secure remote enrolment, while T4.3 refers to the User Continuous Behavioral Authentication (UCBA) technology which will be implemented in order to proceed to continuous assessment of the user through previously collected behavioral patterns. The effectiveness of the technology mentioned on those tasks will be tested throughout the progress of the project through malicious attacks and test that will be carried out.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	167 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

The forementioned features can potentially guarantee that only authorized users have access to the platform, the personal data collected and limits the possibility of any data breach events or unauthorized access or alteration – destruction of the personal data. The level of importance shown by the TANGO project partners on the access monitoring and authentication factor is proven both by the implementation of several different authentication technologies and the dedication of all the partners participating in the related tasks.

T3.2 refers to the development of 3 components that will support and ensure the trustworthiness of the data sharing through the operation of the TANGO platform. Access and usage control of personal data will be supported by the implementation of a data usage control component, the development of a model for calculating a trustworthiness score and the utilization of modules called Ubiquitous Personal Context Vectors (UPCVs). The usage control component will perform privacy risks analysis in terms of data sharing, identifying any possible privacy gaps, based on simulating malicious attacks, while with the trustworthiness score module will keep a score related to the forementioned measurements and the UPCV component will proceed to recommendations in order for the forementioned privacy gaps to be eliminated. Those components will be developed during the course of the project and the level of privacy ensured by them will be dependent to the use cases scenarios - pilots that will be conducted, especially the privacy by design recommendations and the access management and control protocols.

Furthermore, the implementation of additional security measures related to the share of data through the operation of the TANGO platform have been foreseen. The implementation of encryption methods, authentication methods with the help of the forementioned SSI protocol and even the consideration of implementation of anonymization techniques on the personal data adhere to the “privacy by design” development of the TANGO platform. The use of the described in subsection 3.4 self-encryption and decryption methods aim to ensure that only authorized individuals can access the personal data on the cloud server, thus reducing the risk of a data breach event.

The use of the continuous behavioral authentication element has also been assessed. For the purposes of strengthening the access control element, the project partners in charge have proposed the use of an element which may authenticate the user’s device via the online behavioral analysis and learning of the unique patterns and characteristics of a device’s operation such as its hardware configuration, software settings, usage patterns, and network activity.

This element, as described on subsections 4.3 and 4.4, although it could provide a high level of security, needs to be addressed on a personal data perspective. The behavioral analysis of the users will depend also on their personal data, such as the id and characteristics of their device, their ip and mac address, which will be stored and in general processed for the purposes of authentication. Moreover, it can be safely considered that the forementioned personal data will be processed repeatedly, at least every time the user – data subjects uses the TANGO platform.

As a result, the TANGO platform shall inform the users regarding the processing operations of the forementioned personal data during the consent secure process and include them on the personal data that will be described on the consent form, referring in detail to the purposes of such processing operations. Even in the case that the legal basis of legitimate interest may be also suitable in cases as such, since such processing operations aim to strengthen security and protect the interests of both parties, both the data subjects but also the entities providing their services through the use of the TANGO platform, thus assuring the security of transactions, the data subjects shall be informed regarding those processing operations.

Finally, the development of AI technology for the purposes of various factors related to the operation of the TANGO platform has been foreseen in WP5 and our partners have presented the proposed architecture on section 5 of the current deliverable. In terms of privacy, tasks T5.4 and T5.5 refer to the measures implemented via the use of the AI technology in order to achieve data privacy and integrity and ethical compliance. On T5.4, the development of two components is described, the Privacy

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	168 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Enhancing Component (PEC) which aims to provide a tool, based on the NIST Privacy impact assessment methodology, to perform privacy impact assessment and the Privacy Assurance Tool (PAT) which aims to monitor privacy related to the consent and data policy between data subjects and organisations. Those two components will be used as a first step of identification of the implementation of the basic GDPR principles such as the lawfulness of processing, the implementation of security measures according to the GDPR provisions and the anticipation of possible security gaps. As described on the subsection, those components will have a supportive role and should assist the operations of the authorized legal partner / department of the entity who will be using the TANGO platform for their commercial purposes.

To conclude, following the evaluation of the architecture plan proposed on the previous sections, we can safely assume that adequate levels of security are foreseen, especially in terms of access authorization – authentication and data sharing.

9.10 Ethics assessment

Apart from the data privacy assessment, a comprehensive one related to the ethics provisions has also been conducted. In the data management plan of the project the ethics principles have been presented, while instances in which ethical issues may arise are also highlighted. Following this, by the nature of the project, the factors that will be thoroughly examined on this assessment and may cause ethics concerns are the processing of special categories of data and the compliance to the provisions related to the processing operations as a result of a consent provided by the data subjects. Both have been examined earlier on the GDPR perspective and considered compliant to the GDPR provisions but it should be highlighted that the fact that a research is legally permissible according to data protection legislation does not necessarily mean that it will be deemed ethical too.

Regarding the provided consent by the data subjects, as it is highlighted above in section 9.6, the partners provide the necessary guarantee that its provision meets the necessary ethical requirements. The partners of the project have explained / will explain (in case a consent has not been provided yet) to the data subjects in detail the purposes of the project and the processing operation of their data, what their participation in the project will entail and any risks that may be involved. In general, the data subjects have received / will receive all the necessary information detailed in section 9.6 above, thus assuring that there will be no ethical concerns regarding this matter.

Furthermore, it is not expected that the participants - data subject are or will be in any case in a vulnerable position. This characteristic refers to either children / young data subjects or data subjects who may suffer from life threatening diseases or health problems in general. These cases (especially the ones related to health problems) are mostly met in projects related to the development of mechanisms that help the prognosis and / or the treatment of a disease since the participation of patients is needed for the clinical studies. People, especially those suffering from life threatening diseases, are considered to be very vulnerable and not have the mindset in order to proceed to any decisions regarding themselves. It is common for patients suffering from life threatening diseases to accept participation in research projects related to their condition solely because they consider any event / decision that concerns them to be unimportant due to the fact that their life is in danger. As a result, the participation of people as such should be strictly investigated and the provision of their consent for participation and processing of their personal data should be clarified.

In the case of the TANGO project, there are no participants suffering from any life-threatening health condition or children involved since it is not needed for the purposes of the project. The partners acknowledge the fact that the participants should not have those characteristics or age that may cause any ethical concerns in terms of both the consent provided by the data subjects and overall by their participation in general to the project pilots - project.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	169 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Moreover, the processing of special categories of data as stated above will be limited and related only to id authentication mechanisms and access monitoring for the use of the platform on the banking sector and the visa applications. As stated above, such mechanisms are widely used and deemed appropriate for such purposes, taking into account the nature of the transactions. In terms of the visa applications' use cases, the processing of photographs and in general of documents containing facial data of the data subjects - participants is extremely important and necessary in order for the competent authorities to be able to identify the applicant and proceed to the necessary background research. Additionally, the monitoring tools being used on the banking sector's use cases aim to identify the user remotely, on a similar way this procedure is being carried out during the in-person, face to face daily transactions with a bank (id confirmation for example), for the purposes of assuring that the lawful beneficiary has access to the banking records of each data subject.

The re-use of retrospective data could also raise possible ethical issues. The data controllers -partners have ensured that the processing operations on retrospective data are lawful and according to the related framework, most notably the GDPR. Factors such as consent, transparency, privacy and the rights and expectations of the data subjects are key in this case, even if the data is publicly available, and have been taken into consideration.

Finally, the implementation of AI technology may always lead to ethical issues. According to the ethical assessment provided by the "The assessment list for trustworthy Artificial Intelligence (AI TAI)", issued by the High-Level Expert Group on AI, the following factors should be taken into account:

- i. human agency and oversight (level of involvement of the AI on the decision process, related HITL, HOTL or HIC governance mechanisms),
- ii. technical robustness and safety (level of involvement of the AI in case of threats or risks that may arise, of operational accuracy and reliability and possible security measures to anticipate those risks)
- iii. privacy and data governance (whether the use of AI is going to raise issues related to data privacy and the measures implemented to anticipate such issues, privacy "by design"),
- iv. transparency (traceability of the AI operations, explainability of such operations and provision of proper information to the users),
- v. diversity, non-discrimination, and fairness (the AI should correspond to the variety of preferences and abilities in society, according to the purpose of the project, user-centric development and interaction with potential stakeholders)
- vi. environmental and societal impact and
- vii. accountability (obligation to prove ethical and legal compliance)

Following this assessment, we will be able to witness whether the use of AI technology can be deemed lawful, ethical and robust, the three key components, which should be met throughout the system's entire life cycle.

The most crucial factor to be considered is the level of participation of the AI in the decision-making process. Decision bias and/or implementation of wrong decisions as a result of either a mistake on the code or malicious attacks always remain a possibility and the implementation of the proper monitoring tools is crucial. In order to avoid this, on T5.5 the development of the Explainable Artificial Intelligence (XAI) technology is foreseen. Its main features will be to provide machine learning and deep learning model explainability and present the explainability results in an easy-to-understand visual way.

As a result of this technology, explainability reports will be conducted, referring to insights into how the decision-making process of the AI is made. This will provide us with the necessary tools to monitor the decision-making process and avoid any decision-making bias or mistakes that may lead to the violation of the ethical principles, thus assuring that the involvement of the human factor on the decision-making process remains definitive.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	170 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

9.11 Social impact assessment

In terms of social concerns, the goal is to analyze, monitor and manage the intended and unintended social consequences, both positive and negative, of planned interventions (policies, programs, plans, projects) and any social change processes invoked by those interventions. The objective of the social impact assessment and of course of the framework proposed and examined is to ensure that the expected changes maximize the benefits and minimize the costs (both in the present and in the future), especially those related to the people – users and other factors such as the ecosystem.

Taking into consideration that the pilots have not started yet, we cannot safely assume the impact that the use of the TANGO technology may have in the society and it is expected that we will have a clearer view during the pilot phase. In general, four key elements to be examined for the purposes of the social impact assessment are the following:

- i. value created as a result of the use of the framework,
- ii. value experienced by beneficiaries and those affected by it,
- iii. an impact which includes both positive and negative effects and
- iv. an impact marked against the currently followed framework since with TANGO the use of a new, innovative framework is proposed.

Of course, the purpose of this project is to develop a platform which first of all successfully provides services according to each pilot activity, while simultaneously ensuring user-friendly, secure, trustworthy, compliant, fair, transparent, accountable and environmentally sustainable data management. As a result, in order to evaluate its effectiveness, the pilots should be started in order to assess the level of improvement of the provided services in regard to the way this was done previously and compare its function and service to other, currently utilized related, framework. The factors which will be addressed, with the results to be presented on D2.4, are the level that the users will be benefited from in regard to the currently implemented way of services, an estimate of the people benefited by the framework and the way this may happen and the outcome as a result of the use of the TANGO framework. The level of benefit will be defined by assessing the impact of the use of the TANGO framework not only individually to each case – user but to a wider range of stakeholders in order to reach more conclusive results.

Apart from the forementioned, taking into account the purpose of the TANGO project and the proposed framework to be developed during the course of it, we could reach to some considerations, regarding which we could be able to reach to conclusions during the progress of the pilots. For starters, the privacy enhancement components ensure the trustworthiness of the framework in terms of data protection and integrity, with the implementation of the proper technical and privacy by design measures. Ensuring data privacy and in general the importance being shown from the earlier stages of the project regarding this matter, ensures the creation of a “cyber hygienic environment”, a very crucial matter for the users which could only have a positive impact.

Furthermore, the development of components which will have a positive impact on the environment could also ensure the success of the framework. One of the aims of the project is reduction of the environmental footprint generated by the involved infrastructure (data centres), the energy needed to train the AI models that process the collected data and the energy consumed in application level, as described in T5.1-5.3. Such priority, especially in a case of a platform, which for many people does not have a direct relation to the environment (such as for example the use of cars with internal combustion engines or any other functions that directly cause environmental pollution), could also have a positive impact on the users since they could feel that, comparing to other already existent related framework, only by the use of TANGO they could help reduce the environmental pollution.

Finally, the reduction of the human cooperation on some of the use cases – pilots, may also have both positive and negative results socially, especially in terms of the smart hospitality and autonomous

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	171 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

vehicles pilots. The reduction of the human cooperation may lead to faster, easier and more convenient for the user provided services since he/she will be able to proceed to the necessary procedures at their will. On the other hand though, the lack of human cooperation could also potentially have a negative impact for both the users, especially those not familiar to new technologies or those who demand the human presence for the purposes of provision of services, and the organizations who provide the services since human interaction most of the times ensures a better quality for the provided services (for example, in terms of smart hospitality, the hotel may not be able to evaluate the customers in terms of interest and in general state of life and end up mixing families, who most of the time need to have more relaxed vacations, with young adults, who tend to have the opposite lifestyle, thus possibly creating frictions between the customers).

All the forementioned factors will be thoroughly evaluated and assessed during the pilot phase of the project and we will be able to have much more information during the first months of the pilots. For now, we could only refer to a primitive approach of ours to the matter and more information will be available and provided for the purposes of D2.4, following the start of the pilots.

9.12 Considerations

As mentioned above, the reason of conduction of the DIPA+ is to note the processing operations of personal data, recognize any possible risks that may arise regarding the rights and freedoms of the data subjects and evaluate the security measures that have been implemented by the partners and propose solutions in case of insufficient implemented measures. Following the T2.6 guidelines, we proceeded to the necessary actions to carry out the task, receiving feedback from our partners and engaging to workshops, interviews and in general communicating with them in order to collect the necessary information and in general spread awareness related to data protection issues.

On the data processing operations that are taking place until the date of conduction of this assessment and are expected to take place later and mostly during the conduction of the pilots, it is expected that special categories of data will not be included, except from the pilots related to the banking and the visa applications, where it is deemed, as mentioned above, that the processing of such personal data is needed since, first of all, is commonly performed in such occasions and aims to support the development and proper function of the TANGO platform.

On the current circumstances, we can safely assume that the processing and in general project operations are taking place according to the related legal, privacy or ethical provisions, thus eliminating the possibility of any issues arising. The partners have the necessary feedback regarding those matters and the authorized legal partner provides them with the necessary assistance, while simultaneously engaging to various monitoring activities to assure the compliance to the related legal, privacy or ethical provisions and eliminate any privacy concerns.

It should be noted though that there are some ethical principles related to the participation of humans in the project that should be taken into consideration throughout the progress of the project. Human participation should be voluntary and the participants be thoroughly informed about the nature of the project, the details of the research activity and the use of the data that will be collected and consequently processed. No project activity entailing human participation should be conducted without the provision of consent from the respective participant. In addition, clear information about the way a person can withdraw from participation in a given task in any time should be included.

Furthermore, human participants selection criteria should be defined through related tasks, while vulnerable people, children, patients or in general people unable to provide their consent for participating in the project should be excluded. The project partners should duly consider how to integrate relevant considerations by ensuring the balanced representation of women and men in activities of the project

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	172 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

and having regard to gender issues, where relevant and appropriate. Lastly, research activities under the project should overall be carried out by a gender- balanced team of researchers.

It should be noted though that the process of monitoring of the processing operations will be ongoing and the authorized legal partner shall proceed to the monitoring operations described above throughout the progress of the project. As a result, the cooperation and communication with the rest of the partners is of utmost importance and may take place periodically, especially by the time the pilots will start.

9.13 Key legal requirements in terms of privacy and ethics

In order to assure privacy and ethical compliance, the following legislation should be mostly taken into consideration:

1. General Data Protection Regulation (GDPR)
2. European Code of Conduct for Research Integrity
3. “Ethics and data protection” guidelines issued by the European Commission
4. Ethics Guidelines for a Trustworthy AI of the High-Level Expert Group on AI
5. The assessment list for trustworthy Artificial Intelligence (AL TAI) of the High-Level Expert Group on AI

These have been translated into relevant privacy, ethics, legal and societal requirements, as seen below:

Table 67: Privacy, Ethics, Legal, and Societal Requirements (Author: DBC)

ID	Description	Category	Priority
LEG-001	The TANGO system must ensure that the processing operations are taking place lawfully and according to the GDPR provisions.	PRIVACY, ETHICS, LEGAL COMPLIANCE	H
LEG-002	The TANGO system must provide the users with the way that they could exercise their privacy related rights according to the GDPR.	PRIVACY, ETHICS, LEGAL COMPLIANCE	H
LEG-003	The TANGO system must be developed in compliance to the privacy by design related framework in terms of security and data privacy and integrity.	PRIVACY, ETHICS, LEGAL COMPLIANCE	H
LEG-004	The TANGO system must implement the use of the AI with caution and in accordance with the related ethical provisions.	PRIVACY, ETHICS, SOCIAL, LEGAL COMPLIANCE	H
LEG-005	The TANGO system must follow the ethics recommendations regarding the participation of humans on the pilots of the project.	PRIVACY, ETHICS, SOCIAL, LEGAL COMPLIANCE	H

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	173 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

10 Conclusions

TANGO aspires to establish an ecosystem that prioritizes transparency, security, fairness, accountability, and environmental sustainability, adhering to the IDSA & GAIA-X data management and sharing principles for data spaces. Building upon the initial version of user requirements and use case scenarios defined in the previous WP2 deliverable (D2.2), this document represents our initial efforts to translate these concepts into both functional and non-functional requirements. These requirements will serve as a guiding framework for technical solution providers as they embark on implementing or enhancing their offerings to align with the data space paradigm. The elicited technical requirements have been carefully considered in conjunction with the envisaged use case scenarios in the development of the TANGO Reference Architecture which incorporates at its core the critical principles of data sovereignty and trust promoted by the IDSA/GAIA-X reference model and specifications to enable the seamless operation of data spaces and foster robust relationships among participants. An account of the early assessment of TANGO's Privacy, Ethical, Social, and Legal Impact has been also included in this deliverable. Alongside this preliminary assessment, we have outlined a comprehensive set of technical and operational measures designed to safeguard the fundamental rights to security and privacy of data subjects.

Deliverable D2.3 offers an insightful overview of three pivotal tasks within the project, which will continue to unfold until its final year of implementation. This report provides a comprehensive snapshot of the proposed framework's design and development, laying a strong foundation for the subsequent WPs (WP3-WP6) that will produce the TANGO framework prototype. The findings presented in D2.3 will actively contribute to the ongoing development and future integration efforts. TANGO's requirements, reference architecture, and impacts will undergo periodic review and refinement to ensure the framework's alignment with the dynamic relevant technological and legal landscapes. The conclusive outcomes of these activities will be documented in the last deliverable of WP2, namely D2.4 (M28), taking into account essential feedback collected from the development and integration processes.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	174 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

11 Bibliography

- [1] TANGO Consortium, “D2.2: User Needs and Requirements & Use Case Scenarios,” 2023.
- [2] TANGO Consortium, “D2.1 State-of-the-Art & GAP analysis Distributed Data Management,” 2023.
- [3] K. Brennan, A Guide to the Business Analysis Body of Knowledge, 2009.
- [4] ISO/IEC JTC 1/SC 7, “ISO/IEC 25010:2011 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models,” 2011.
- [5] K. M. Adams, Nonfunctional requirements in systems analysis and design, Springer international publishing, 2015.
- [6] N. Saher, F. Baharom and R. Romli, “Identification of Sustainability Characteristics and Sub-Characteristics as Non-Functional Requirement for Requirement Change Management in Agile,” *International Journal of Scientific & Technology Research*, 2020.
- [7] G. A. García-Mireles and H. A. Villa-Martínez, “Practices for addressing environmental sustainability through requirements processes,” in *Trends and Applications in Software Engineering: Proceedings of the 6th International Conference on Software Process Improvement*, 2018.
- [8] P. Homan, “What IT architects can do to design for sustainability,” 30 January 2023. [Online]. Available: <https://www.redhat.com/architect/sustainability-nfr>.
- [9] J. Frankenfield, “51% Attack: Definition, Who Is At Risk, Example, and Cost (investopedia.com),” Investopedia, 2023.
- [10] Dworkin et Al., “SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions,” 2015.
- [11] Cruz et Al., “Keccak: The New SHA-3 Encryption Standard,” 2013.
- [12] O’Whielacronx et Al., “Introducing BLAKE2 – an alternative to SHA-3, SHA-2 and MD5,” 2012.
- [13] “Hyperledger ARIES,” 12 09 2023. [Online]. Available: <https://gitlab.com/GAIA-X/data-infrastructure-federation-services/pcm>.
- [14] “DIDCOMM protocol,” [Online]. Available: <https://identity.foundation/didcomm-messaging/spec/>. [Accessed 12 09 2023].
- [15] e. a. O. Terbu, “OpenID Connect for Verifiable Presentations,” April 2023. [Online]. Available: https://openid.net/specs/openid-4-verifiable-presentations-1_0.html. [Accessed 09 12 2023].
- [16] e. a. K. Yasuda, “Self-Issued OpenID Provider,” 01 01 2023. [Online]. Available: https://openid.net/specs/openid-connect-self-issued-v2-1_0.html. [Accessed 09 12 2023].
- [17] Data Spaces Business Alliance, “Technical Convergence-Discussion Document v. 1.0.1,” Data Spaces Business Alliance, 2022.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	175 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

- [18] GXFS, “The Eclipse Foundation Takes Over GAIA-X Federation Services: More Cross-Linking for Communities for Digital Sovereignty,” 2023.
- [19] V. Bolgouras et al., “Trusted and Secure Self-Sovereign Identity framework,” New York, 2022.
- [20] Nicholas Dominic et al., “Society with Trust: A Scientometrics Review of Zero- knowledge Proof Advanced Applications in Preserving Digital Privacy for Society 5.0,” 2022.
- [21] L. Masure, N. Belleville, E. Cagli, M.-A. Cornélie, D. Couroussé, C. Dumas and L. Maingault, “Deep learning side-channel analysis on large-scale traces: A Case Study on a Polymorphic AES,” in *European Symposium on Research in Computer Security*, 2020.
- [22] “Tornado VM,” [Online]. Available: <https://www.tornadovm.org>. [Accessed 2023].
- [23] A. Stratikopoulos, “Migrating existing TornadoVM applications to TornadoVM v0.15,” 13 February 2023. [Online]. Available: <https://www.tornadovm.org/post/migrating-existing-tornadovm-applications-to-tornadovm-v0-15>.
- [24] H. Rong, H. Zhang, S. Xiao, C. Li and C. Hu, “Optimizing energy consumption for data centers,” *Renewable and Sustainable Energy Reviews*, vol. 58, pp. 674--691, may 2016.
- [25] E. Oró, V. Depoorter, A. Garcia and J. Salom, “Energy efficiency and renewable energy integration in data centres. Strategies and modelling review,” *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 429--445, feb 2015.
- [26] European Commission and Joint Research Centre, P. Bertoldi, M. Avgerinou and L. Castellazzi, Trends in data centre energy consumption under the European code of conduct for data centre energy efficiency, Publications Office, 2017.
- [27] B. L. Christian , “In the data center, power and cooling costs more than the it equipment it supports,” 1 February 2007. [Online]. Available: <https://www.electronics-cooling.com/2007/02/in-the-data-center-power-and-cooling-costs-more-than-the-it-equipment-it-supports/>.
- [28] A. Greenberg, J. Hamilton, D. A. Maltz and P. Patel, “The cost of a cloud,” *SIGCOMM Computer Communication Review*, vol. 39, no. 1, pp. 68--73, dec 2008.
- [29] IRENA, “Renewable Power Generation Costs in 2021,” International Renewable Energy Agency, Abu Dhabi, 2022.
- [30] D. Patterson, J. Gonzalez, U. Hölzle, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. R. So, M. Texier and J. Dean, “The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink,” *Computer*, vol. 55, p. 18–28, jul 2022.
- [31] [Online]. Available: <https://transparency.entsoe.eu/>.
- [32] Internet Security Research Group (ISRG), “Let's Encrypt,” [Online]. Available: <https://letsencrypt.org>.
- [33] RegisFrey, “File:MVC-process.svg,” RegisFrey, Public domain, via Wikimedia Commons, May 2010. [Online]. Available: <https://commons.wikimedia.org/w/index.php?curid=10298177>. [Accessed 25 July 2023].
- [34] International Data Spaces Association, “Data Usage Control in the International Data Spaces 3.0,” [Online]. Available: <https://internationaldataspaces.org/data-sovereignty-updated-position-paper-on-data-usage-control-in-the-ids/>.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	176 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

- [35] Eitel Andreas et Al., “Usage Control in the International Dataspaces,” IDS, 2021.
- [36] “Github,” [Online]. Available: <https://github.com/about>. [Accessed 25 09 2023].
- [37] International Data Spaces Association, “IDSA White Paper Specification IDS Clearing House,” [Online]. Available: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-White-Paper-Specification-IDS-Clearing-House-.pdf. [Accessed 26 09 2023].
- [38] W3C Recommendation, “Decentralized Identifiers (DIDs) v1.0,” 2022. [Online]. Available: <https://www.w3.org/TR/did-core/#dfn-verifiable-data-registries>. [Accessed 29 09 2023].
- [39] “Verifiable Credentials Framework,” EBSI, [Online]. Available: <https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/EBSI+Verifiable+Credentials>. [Accessed 28 09 2023].
- [40] EBSI, “EBSI DID Method,” [Online]. Available: <https://ec.europa.eu/digital-building-blocks/wikis/display/EBSIDOC/EBSI+DID+Method>. [Accessed 28 09 2023].
- [41] Michael Prorock et Al., “did:web Method Specification,” [Online]. Available: <https://w3c-ccg.github.io/did-method-web/>. [Accessed 28 09 2023].
- [42] Manu Sporny et Al., “The did:key Method v0.7,” [Online]. Available: <https://w3c-ccg.github.io/did-method-key/>. [Accessed 28 09 2023].
- [43] “Docker registry,” [Online]. Available: <https://docs.docker.com/registry/>. [Accessed 28 09 2023].
- [44] “Jfrog artifactory,” [Online]. Available: <https://jfrog.com/artifactory/>. [Accessed 28 09 2023].
- [45] Otto B. et al., “GAIA-X and IDS,” International Data Spaces Association , Dortmund, 2021.
- [46] Sebastian Steinbuss et al., “Reference Architecture Model - Version 3.0,” Berlin, 2019.
- [47] Giulia Giussani, Sebastian Steinbuss, “Data Connector Report,” International Data Spaces Association, Dortmund, Germany, 2023.
- [48] “IDSA certification,” [Online]. Available: <https://internationaldataspaces.org/offers/certification/>. [Accessed 28 09 2023].
- [49] “IDS reference test suite,” [Online]. Available: <https://docs.internationaldataspaces.org/knowledge-base/ids-testbed-1>. [Accessed 28 09 2023].
- [50] GAIA-X, “GAIA-X labelling criteria,” [Online]. Available: https://GAIA-X.eu/wp-content/uploads/2022/05/GAIA-X-labelling-criteria-v22.04_Final.pdf. [Accessed 28 09 2023].
- [51] “GAIA-X Digital clearing house,” [Online]. Available: <https://GAIA-X.eu/gxdch/>. [Accessed 28 09 2023].
- [52] “GDPR-Info,” 27 April 2016. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>. [Accessed 29 09 2023].
- [53] “Truffle Language Implementation Framework,” [Online]. Available: <https://www.graalvm.org/latest/graalvm-as-a-platform/language-implementation-framework/>.
- [54] “Verifiable Credentials Data Model v1.1,” [Online]. Available: <https://www.w3.org/TR/vc-data-model/>. [Accessed 12 09 2023].
- [55] “W3C Recommendation Decentralized Identifiers (DIDs) v1.0,” [Online]. Available: <https://www.w3.org/TR/2022/REC-did-core-20220719/>. [Accessed 12 09 2023].

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	177 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

[56] “W3C Working Draft, Verifiable Credentials Data Model v2.0,” 12 09 2023. [Online]. Available: <https://www.w3.org/TR/vc-data-model-2.0/>.

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment				Page:	178 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Annexes

* N/A fields have been eliminated from the tables to conserve space.

A. System Requirements for Distributed Privacy-preserving Data Management and Storage

A.1 Functional Requirements

Table 68: Blockchain-based Data Storage and Sharing: FR-DATA-BDSS-001

ID	FR-DATA-BDSS-001	Author	NOR
Name	Explicit customer consent management		
Description	The TANGO system must be able to collect and manage customer consent for every act of customer data sharing. TANGO system must be able to keep the collected consent, enforce consented policies and demonstrate it to the auditing authority.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-TDS-PA-010, UR-GCO-PA-011		
Rationale	This feature ensures legal compliance of the data sharing component as collecting and correctly managing customer consent is required by GDPR.		

Table 69: Blockchain-based Data Storage and Sharing: FR-DATA-BDSS-002

ID	FR-DATA-BDSS-002	Author	NOR
Name	Private data sharing in regulated environments		
Description	The TANGO system must ensure that the data sharing mechanism is fully compliant with the relevant regulations concerning private data handling.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-TDS-PA-010, UR-GCO-PA-011		
Rationale	This feature ensures compliance of the data sharing component to GDPR requirements.		

Table 70: Blockchain-based Data Storage and Sharing: FR-DATA-BDSS-003

ID	FR-DATA-BDSS-003	Author	NOR
Name	Immutable audit logs		
Description	The TANGO system must ensure that the data access logs are stored immutably, in a safe and legally compliant manner, and are available to auditing authorities upon request.		
Priority	HIGH		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	179 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Technology Category	DATA
Related User Requirement	UR-PCY-PA-012, UR-PCY-PA-013
Rationale	This feature enables automated compliance audit, as from the immutable audit logs it is possible to demonstrate that all necessary data checks were performed, and data accesses were consented by the data subjects.

Table 71: Trustworthy Data Sharing: FR-DATA-TDS-001

ID	FR-DATA-TDS-001	Author	FHG
Name	Privacy Risk estimation		
Description	Simulated privacy attacks based on predefined adversary settings against the data sharing process or the shared data to estimate the risk of possible privacy leakages. The privacy risk analysis must be performed to aid in the creation of a suitable usage control policy that is then enforced later.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-GCO-HT-011, UR-PCY-HT-012, UR-TDS-HT-010, UR-DUP-AV-005, UR-GCO-AV-006, UR-DAR-M1-011, UR-DAR-M1-012, UR-GCO-M1-013, UR-GCO-M2-007, UR-GCO-M2-007, UR-TUA-M2-016, UR-DMN-M2-018		
Rationale	To share or process personal data according to the GDPR user consent must be given. However, many uses cases evolve over time requiring that data is shared without the explicit user consent. Applying privacy preserving technologies make it possible to do that. For that privacy preserving technologies like anonymization or Secure Multiparty computation can be used. Nevertheless, this only works if the data is protected to a sufficient degree which cannot be done since the threat level is unknown without a prior analysis. To understand which privacy enhancing technology should be used and to what level, a privacy risk analysis must be performed. Subsequently the needed privacy guarantee can be enforced by a usage control component.		

Table 72: Trustworthy Data Sharing: FR-DATA-TDS-002

ID	FR-DATA-TDS-002	Author	FUJ_DE
Name	Scoring the Trustworthiness of the vehicle and the CAVride system.		
Description	The TANGO system should facilitate the assessment of the vehicles' and the CAVride's system regarding Security, Privacy, Reliability, Resilience and Safety. Based on the assessment (attributes and degree of fulfilment) provided by the use case, a weighted scoring enables evaluation of Trustworthiness and mitigation measures depending on the Trustworthiness scores. Thus, the TANGO system must provide a module configuring the Trustworthiness calculation and a module calculating the Trustworthiness scores based on this configuration.		
Priority	HIGH		
Technology Category	DATA		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	180 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Related User Requirement	UR-DIN-AV-002, UR-DIN-AV-003, UR-DIN-AV-004, UR-DIN-AV-006, UR-DIN-AV-010, UR-DIN-AV-012
Rationale	<ol style="list-style-type: none"> 1) The communication between user, CAVride system and vehicle needs to ensure data integrity and prevention or at least detection of potential cyber-attacks. 2) Each communicating system should continuously be assessed on conformity ensuring its Trustworthiness. 3) The scoring module depends on the assessment within each system; it probably should be deployed locally on the use case side.

Table 73: Trustworthy Data Sharing: FR-DATA-TDS-003

ID	FR-DATA-TDS-003	Author	FTS DE
Name	Scoring the Trustworthiness of the Data Gate Keeper system.		
Description	The TANGO system should facilitate the assessment of the Data Gate Keeper system regarding Security, Privacy, Reliability, Resilience and Safety. Based on the assessment (attributes and degree of fulfilment) provided by the use case, a weighted scoring enables evaluation of Trustworthiness and mitigation measures depending on the Trustworthiness scores. Thus, the TANGO system must provide a module configuring the Trustworthiness calculation and a module calculating the Trustworthiness scores based on this configuration.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-TUA-M2-011, UR-PCY-M2-006, UR-PMT-M2-014		
Rationale	The Data Gate Keeper system shall be the gateway to the use case systems and data. Thus, it is highly endangered to be compromised by an attacker and its Trustworthiness shall be ensured.		

Table 74: Trustworthy Data Sharing: FR-DATA-TDS-004

ID	FR-DATA-TDS-004	Author	VTT
Name	Collaborative recommendations for hotel guests		
Description	<p>The TANGO system should provide a data repository of few kilobytes for each end guest and each vendor (i.e. hotel, activity provider etc...).</p> <ol style="list-style-type: none"> a) Other tasks in TANGO should preferably provide UI/UX development for guest and vendor frontends, for uniform user experience within TANGO system. 		
Priority	MEDIUM		
Technology Category	DATA		
Related User Requirement	UR-DMN-HT-008, UR-UIN-HT-014, UR-UIN-HT-015, UR-DMN-HT-018		
Rationale	<ol style="list-style-type: none"> 1) Hotel guests will get a personalized view of the options they have during their stay, and throughout their stay. This may mean sorting the offerings by assumed relevance, as predicted by the recommendation engine. 		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	181 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	<p>2) The guest data upon which the personalized view of based on will be stored on personal wallets.</p> <p>3) Vendors may co-operate by sharing item specific VRN collections, without compromising users' privacy.</p> <ul style="list-style-type: none"> - This is actually a parallel with Amazon platform recommendations (people who bought this also bought that), in which a single player (i.e. Amazon itself) provides and monetizes recommendations based on user behaviour on their single platform. - Recommendations for users are created in a multi-vendor environment while respecting privacy, by exchanging data that is not private by nature.
--	---

Table 75: Trustworthy Data Sharing: FR-DATA-TDS-005

ID	FR-DATA-TDS-005	Author	VTT
Name	Privacy-preserving business intelligence		
Description	The TANGO system should provide user transaction data, e.g., shopping baskets.		
Priority	MEDIUM		
Technology Category	DATA		
Related User Requirement	UR-DMN-RT-004, UR-DMN-RT-007, UR-DMN-RT-008, UR-DMN-RT-009, UR-TDS-RT-012, UR-DAR-RT-014, UR-GCO-RT-017, UR-DMN-RT-022		
Rationale	Provide business intelligence, such as item clustering (including people-who-bought-this-also-bought-that) and defining user profiles for subsequent analysis using inherently privacy-enhancing technology.		

Table 76: Confidentiality and Privacy by Design: FR-DATA-CPD-001

ID	FR-DATA-CPD-001	Author	UMU
Name	Policy Definition		
Description	The TANGO system must allow an administrator or policy authors to define the policies based on attributes.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-DMN-RT-005, UR-DMN-RT-006, UR-DMN-RT-007, UR-DMN-RT-008, UR-DMN-RT-009, UR-DMN-RT-010, UR-DMN-RT-011, UR-TDS-RT-012, UR-TUA-M2-001, UR-DAC-M1-008, UR-DST-M1-005, UR-DAR-M1-009, UR-TUA-M2-002		
Rationale	This involves specifying the conditions and constraints that must be satisfied for a policy to be enforced. Policies may include attributes related to user roles, resource properties, environmental factors, or any other relevant context.		

Table 77: Confidentiality and Privacy by Design: FR-DATA-CPD-002

ID	FR-DATA-CPD-002	Author	UMU
----	-----------------	--------	-----

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	182 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Name	Access Control
Description	The TANGO system must enforce access control decisions based on policies evaluation.
Priority	HIGH
Technology Category	DATA
Related User Requirement	UR-DMN-RT-005, UR-DMN-RT-006, UR-DMN-RT-007, UR-DMN-RT-008, UR-DMN-RT-009, UR-DMN-RT-010, UR-DMN-RT-011, UR-TDS-RT-012, UR-TUA-M2-001, UR-DAC-M1-008, UR-DST-M1-005, UR-DAR-M1-009, UR-TUA-M2-002
Rationale	This includes granting or denying access to resources based on the attributes of the requesting user and the attributes associated with the resource being accessed.

Table 78: Confidentiality and Privacy by Design: FR-DATA-CPD-003

ID	FR-DATA-CPD-003	Author	UMU
Name	Granular Policy Control		
Description	The TANGO system must apply fine-grained control over data access and usage.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-DMN-RT-005, UR-DMN-RT-006, UR-DMN-RT-007, UR-DMN-RT-008, UR-DMN-RT-009, UR-DMN-RT-010, UR-DMN-RT-011, UR-TDS-RT-012, UR-TUA-M2-001, UR-DAC-M1-008, UR-DST-M1-005, UR-DAR-M1-009, UR-TUA-M2-002, UR-DMN-M2-017, UR-PMT-M1-021, UR-DUP-PA-007, UR-DUP-PA-008		
Rationale	They should specify who can access the data, under what conditions, and for what purposes. This includes defining roles and permissions, specifying conditions for access, and enforcing restrictions on data usage.		

Table 79: Confidentiality and Privacy by Design: FR-DATA-CPD-004

ID	FR-DATA-CPD-004	Author	UMU
Name	Data Lifecycle Management		
Description	The TANGO system should define the appropriate security controls at each stage of the data lifecycle.		
Priority	MEDIUM		
Technology Category	DATA		
Related User Requirement	UR-DMN-RT-005, UR-DMN-RT-006, UR-DMN-RT-007, UR-DMN-RT-008, UR-DMN-RT-009, UR-DMN-RT-010, UR-DMN-RT-011, UR-TDS-RT-012, UR-TUA-M2-001, UR-DAC-M1-008, UR-DST-M1-005, UR-DAR-M1-009, UR-TUA-M2-002		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	183 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Rationale	Sticky policies must address the entire lifecycle of data, including creation, storage, transmission, processing, sharing, and destruction. Policies should define the appropriate security measures and controls at each stage of the data lifecycle.
-----------	--

Table 80: Confidentiality and Privacy by Design: FR-DATA-CPD-005

ID	FR-DATA-CPD-005	Author	UMU
Name	Policy Enforcement		
Description	The TANGO system must involve mechanisms to enforce policies and restrictions.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-DMN-RT-005, UR-DMN-RT-006, UR-DMN-RT-007, UR-DMN-RT-008, UR-DMN-RT-009, UR-DMN-RT-010, UR-DMN-RT-011, UR-TDS-RT-012, UR-TUA-M2-001, UR-DAC-M1-008, UR-DST-M1-005, UR-DAR-M1-009, UR-TUA-M2-002		
Rationale	Sticky policies require mechanisms to enforce policy rules and restrictions. This may involve integrating with existing systems or applications to enforce access controls, encryption, data loss prevention measures, or other security mechanisms.		

Table 81: Confidentiality and Privacy by Design: FR-DATA-CPD-006

ID	FR-DATA-CPD-006	Author	UMU
Name	Policy-Based Encryption/Decryption		
Description	The TANGO system must support the enforcement of defined policies by an encryption/decryption mechanism.		
Priority	HIGH		
Technology Category	DATA		
Related User Requirement	UR-DMN-RT-005, UR-DMN-RT-006, UR-DMN-RT-007, UR-DMN-RT-008, UR-DMN-RT-009, UR-DMN-RT-010, UR-DMN-RT-011, UR-TDS-RT-012, UR-TUA-M2-001, UR-DAC-M1-008, UR-DST-M1-005, UR-DAR-M1-009, UR-TUA-M2-002		
Rationale	It must be able to encrypt data based on the defined policies and decrypt it only for authorized users or entities.		

Table 82: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: FR-DATA-SED-T-001

ID	FR-DATA-SED-T-001	Author	DUT
Name	Secure Data Encryption		
Description	The TANGO system must implement robust data encryption techniques to protect sensitive information generated and collected by autonomous vehicles or smart manufacturing systems. Encryption ensures that data		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	184 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	remains unreadable to unauthorized parties and maintains its confidentiality throughout storage and transmission.
Priority	HIGH
Technology Category	DATA
Related User Requirement	UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-AV-009, UR-PCY-AV-012, UR-PCY-M1-014, UR-PCY-RT-018, UR-PCY-RT-019
Rationale	This requirement is essential to safeguard the integrity and privacy of the data. By encrypting the data, the TANGO system guarantees that even if unauthorized individuals gain access to the data, they will be unable to decipher its contents, ensuring compliance with data security and privacy principles.

Table 83: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: FR-DATA-SEDT-002

ID	FR-DATA-SEDT-002	Author	DUT
Name	Multi-Factor Authentication		
Description	The TANGO system should support multi-factor authentication mechanisms, requiring users to provide multiple authentication factors such as passwords, biometrics, or physical tokens to access sensitive data or system functionalities. This adds an extra layer of security, ensuring that only authorized individuals with valid credentials can gain access.		
Priority	MEDIUM		
Technology Category	DATA		
Related User Requirement	UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-AV-009, UR-PCY-AV-012, UR-PCY-M1-014, UR-PCY-RT-018, UR-PCY-RT-019		
Rationale	Multi-factor authentication enhances the overall security of the system by verifying the identity of users through multiple independent factors. By incorporating this requirement, the TANGO system significantly reduces the risk of unauthorized access, data breaches, and potential misuse, aligning with data security and privacy principles.		

Table 84: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: FR-DATA-SEDT-003

ID	FR-DATA-SEDT-003	Author	DUT
Name	Highly Efficient Data Encryption		
Description	The TANGO system should implement encryption algorithms and techniques that are highly efficient in terms of computational resources and processing time. The encryption process should be optimized to minimize any performance impact on the overall system functionality.		
Priority	MEDIUM		
Technology Category	DATA		
Related User Requirement	UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-AV-009, UR-PCY-AV-012, UR-PCY-M1-014, UR-PCY-RT-018, UR-PCY-RT-019		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	185 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Rationale	This requirement addresses the need for efficient encryption processes within the TANGO system. By ensuring high efficiency, the system can perform encryption operations rapidly and seamlessly, minimizing any potential disruptions to real-time data processing, transmission, or system performance. Efficient encryption is crucial for maintaining the overall system responsiveness and meeting the operational requirements of use cases.
-----------	--

A.2 Non-Functional Requirements

Table 85: Blockchain-based Data Storage and Sharing: NFR-PER-BDSS-001

ID	NFR-PER-BDSS-001	Author	NOR
Name	Performance efficiency		
Description	The TANGO system must deliver reasonable performance		
Priority	HIGH		
NFR Category	PERFORMANCE / USABILITY		
Rationale	Human-faced user interface should be responsive enough to deliver satisfactory user experience. Therefore, a response time of an API call to data sharing component should never exceed 30 seconds, and for the best user experience it should be less than 5 seconds.		

Table 86: Blockchain-based Data Storage and Sharing: NFR-REL-BDSS-001

ID	NFR-REL-BDSS-001	Author	NOR
Name	Guaranteed data delivery		
Description	The TANGO system must guarantee 100% data delivery when data sharing component is involved.		
Priority	HIGH		
NFR Category	RELIABILITY		
Related User Requirement	UR-PCY-PA-012, UR-PCY-PA-013		
Rationale	When data enters the system, the user must be ensured that the recipient will receive the data eventually (depending on external circumstances).		

Table 87: Blockchain-based Data Storage and Sharing: NFR-SEC-BDSS-001

ID	NFR-SEC-BDSS-001	Author	NOR
Name	Secure data handling		
Description	The TANGO system must guarantee that user data is being handled securely.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-PCY-PA-012, UR-PCY-PA-013		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	186 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Rationale	Security and privacy must be preserved and should be of the highest priority every time we are dealing with customer's private data.
-----------	--

Table 88: Trustworthy Data Sharing: NFR-SEC-TDS-001

ID	NFR-SEC-TDS-001	Author	FHG
Name	Privacy Assurance		
Description	The TANGO system could provide the continuous Privacy Risk Score to a monitoring system, which could visualize the Privacy for easy evaluation.		
Priority	LOW		
NFR Category	SECURITY / PRIVACY		
Rationale	The monitoring of the Privacy scores could be used for visualization or as input for a SIEM protecting the cyber-security of the overall system.		

Table 89: Trustworthy Data Sharing: NFR-SEC-TDS-002

ID	NFR-SEC-TDS-002	Author	FUJ_DE
Name	Trustworthiness Assurance		
Description	The TANGO system could provide the continuous Trustworthiness Score to a monitoring system, which could visualize the Trustworthiness of the vehicle or the platform for easy evaluation by a user.		
Priority	LOW		
NFR Category	SECURITY / USABILITY		
Related User Requirement	UR-DIN-AV-002, UR-DIN-AV-003, UR-DIN-AV-004, UR-DIN-AV-006, UR-TUA-M2-012, UR-TUA-M2-013		
Rationale	The monitoring of the Trustworthiness scores could be used for visualization in the vehicle or as input for a SIEM protecting the cyber-security of the overall system. The visualization of the Trustworthiness Score in the vehicles' screen may even assure Trust of the passenger.		

Table 90: Trustworthy Data Sharing: NFR-SEC-TDS-003

ID	NFR-SEC-TDS-003	Author	FUJ_DE
Name	Trustworthiness Assurance		
Description	The TANGO system could facilitate the assessment of the regular operations of the shopfloor.		
Priority	LOW		
NFR Category	SECURITY / USABILITY		
Related User Requirement	UR-PCY-M2-015		
Rationale	The monitoring of the Trustworthiness scores could be used for monitoring the regular operation on the shopfloor. This way anomalies originating from cyber-attacks as well as from false updates or maintenance can be detected early.		

Table 91: Trustworthy Data Sharing: NFR-UI-TDS-001

ID	NFR-UI-TDS-001	Author	VTT
----	----------------	--------	-----

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	187 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Name	List of alternative activities in order of personal predicted relevance
Description	The TANGO system should provide hotel guests recommendations in a user-friendly personalized order.
Priority	MEDIUM
NFR Category	USABILITY
Related User Requirement	UR-DMN-HT-006, UR-TDS-HT-010
Rationale	Hotel guests do not have to go through the complete list of alternative activities in order to find ones that suite best their personal needs.

Table 92: Trustworthy Data Sharing: NFR-UI-TDS-002

ID	NFR-UI-TDS-002	Author	VTT
Name	Easy business analytics		
Description	The TANGO system should provide business intelligence from retail locations that are nor per se allowed to share customer data with each other.		
Priority	MEDIUM		
NFR Category	USABILITY		
Related User Requirement	UR-DMN-RT-005, UR-DMN-RT-008, UR-DAR-RT-015		
Rationale	As a result, retail professionals get an easy analysis of shopping baskets and predicted behaviour of most typical user groups. The analysis may be in a form of a spreadsheet, to be visualized and further analyzed in tools like Excel.		

Table 93: Confidentiality and Privacy by Design: NFR-SEC-CPD-001

ID	NFR-SEC-CPD-001	Author	UMU
Name	Confidentiality and Integrity insurance		
Description	The TANGO system must prioritize and enforce robust security measures to protect data.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-PCY-RT-018, UR-PCY-PA-012, UR-PCY-PA-013, UR-PCY-M2-006, UR-DAC-M1-007, UR-TDS-M1-002, UR-DUP-AV-005, UR-PCY-AV-012, UR-PCY-AV-004, UR-PCY-AV-003, UR-DIN-AV-002, UR-TDS-HT-019, UR-PCY-HT-012, UR-TDS-HT-010, UR-DMN-M2-017		
Rationale	Modules must prioritize authentication and confidentiality to protect sensitive data from unauthorized access, tampering or breaches. This can include encryption, access controls and secure storage.		

Table 94: Confidentiality and Privacy by Design: NFR-SEC-CPD-002

ID	NFR-SEC-CPD-002	Author	UMU
Name	Privacy Compliance		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	188 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Description	The TANGO system must ensure the privacy of individuals and sensitive data aligned with GDPR.
Priority	HIGH
NFR Category	SECURITY
Related User Requirement	UR-GCO-RT-017, UR-GCO-M1-013, UR-GCO-PA-011, UR-GCO-M2-007, UR-GCO-HT-011, UR-GCO-AV-006
Rationale	Definition of appropriate restrictions and controls on data usage and sharing aligned with GDPR to ensure the privacy of sensitive information.

Table 95: Confidentiality and Privacy by Design: NFR-PER-CPD-001

ID	NFR-PER-CPD-001	Author	UMU
Name	Scalability and responsiveness		
Description	The TANGO system should not introduce significant performance overhead that could impact system responsiveness.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		
Related User Requirement	UR-DMN-HT-005, UR-DMN-HT-008, UR-DMN-HT-013, UR-DIN-AV-002, UR-TDS-M1-002, UR-DUP-M1-006, UR-PCY-M2-015, UR-DMN-RT-002		
Rationale	Access and usage control should be scalable to handle growing volumes of data, users and policy rules. It should not mean a significant decrease in efficiency or performance, allowing real-time decision-making and minimal latency.		

Table 96: Confidentiality and Privacy by Design: NFR-REL-CPD-001

ID	NFR-REL-CPD-001	Author	UMU
Name	Accuracy and system availability		
Description	The TANGO system must ensure continuous enforcement of policies.		
Priority	HIGH		
NFR Category	RELIABILITY		
Related User Requirement	UR-DMN-HT-005, UR-DMN-HT-008, UR-DMN-HT-013, UR-DIN-AV-002, UR-TDS-M1-002, UR-DUP-M1-006, UR-PCY-M2-015, UR-DMN-RT-002		
Rationale	The whole process of access control (from policy definition to usage control) must be reliable and available to ensure continuous enforcement of data handling rules.		

Table 97: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: NFR-MAIN-SED-T-001

ID	NFR-MAIN-SED-T-001	Author	DUT
Name	Modularity of encryption components		
Description	The TANGO system must be designed with a modular architecture, where the self-encryption and decryption component is a self-contained and easily maintainable module. It should be capable of independent updates or enhancements without affecting the overall system functionality.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	189 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Priority	HIGH
NFR Category	MAINTABILITY
Related User Requirement	UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-AV-009, UR-PCY-AV-012, UR-PCY-M1-014, UR-PCY-RT-018, UR-PCY-RT-019
Rationale	This requirement focuses on the maintainability of the self-encryption and decryption component. By ensuring modularity, the component can be easily maintained, enhanced, or replaced as needed, minimizing the effort and impact on the rest of the system. Modularity promotes code reusability, scalability, and facilitates efficient bug fixes or feature updates, ensuring long-term maintainability and adaptability of the TANGO system.

Table 98: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: NFR-PER-SEDT-001

ID	NFR-PER-SEDT-001	Author	DUT
Name	Performance Efficiency		
Description	The TANGO system should exhibit high-performance efficiency in terms of encryption, decryption, and multi-factor information recovery operations. The component should be optimized to minimize processing time, computational resources, and memory usage, without compromising the overall system performance.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		
Related User Requirement	UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-AV-009, UR-PCY-AV-012, UR-PCY-M1-014, UR-PCY-RT-018, UR-PCY-RT-019		
Rationale	This requirement addresses the need for efficient and effective performance of the self-encryption and decryption component. By optimizing performance efficiency, the component ensures that encryption, decryption, and multi-factor information recovery operations are executed promptly, reducing any potential delays in data processing or system responsiveness. This is vital for real-time data generation, transmission, and processing in autonomous vehicles or smart manufacturing systems.		

Table 99: Self-encryption and Decryption Techniques with Multi-Factor Information Recovery Mechanisms: NFR-COMP-SEDT-001

ID	NFR-COMP-SEDT-001	Author	DUT
Name	Interoperability		
Description	The TANGO system should demonstrate interoperability with commonly used communication protocols and data formats in the autonomous vehicle or smart manufacturing domain as far as self-encryption and decryption techniques are considered. It should seamlessly exchange information with other systems to facilitate smooth integration and collaboration.		
Priority	MEDIUM		
NFR Category	COMPATIBILITY		
Related User Requirement	UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-AV-009, UR-PCY-AV-012, UR-PCY-M1-014, UR-PCY-RT-018, UR-PCY-RT-019		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	190 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Rationale	Including this requirement ensures that the self encryption and decryption component can effectively communicate and exchange data with other components and systems in the ecosystem. Interoperability allows for seamless information flow, enabling efficient data processing, analysis, and decision-making across different systems. This is crucial for the successful integration and operation of the TANGO system within complex technological environments.
-----------	---

B. System Requirements for Distributed Trust Management Framework

B.1 Functional Requirements

Table 100: Self-sovereign identity management: FR-TRUST-SSIM-001

ID	FR-TRUST-SSIM-001	Author	UMU
Name	Key Pair generation		
Description	The TANGO dp-ABC cryptographic system must generate a public and a private key.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TDS-HT-010, UR-DAC-AV-007, UR-TUA-AV-001, UR-DAC-M1-008, UR-DAC-M1-007, UR-DAC-M1-004, UR-TUA-M1-003, UR-TDS-M1-002, UR-TDS-M1-002, UR-TUA-M1-003, UR-DAC-M1-004, UR-DAC-M1-007, UR-DAC-M1-008, UR-TUA-PA-001, UR-TUA-PA-002 , UR-TUA-PA-003, UR-TUA-PA-004 , UR-TDS-PA-010, UR-TUA-RT-001, UR-TDS-RT-012		
Rationale	Creation of the pair keys will address confidentiality- data can only be decrypted using the private key only known by the owner, and authenticity- public key used to verify digital signature created with the private key.		

Table 101: Self-sovereign identity management: FR-TRUST-SSIM-002

ID	FR-TRUST-SSIM-002	Author	UMU
Name	Signature generation		
Description	The TANGO dp-ABC cryptographic system must be able to generate a digital signature.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TDS-HT-010, UR-DAC-AV-007, UR-TUA-AV-001, UR-DAC-M1-008, UR-DAC-M1-007, UR-DAC-M1-004, UR-TUA-M1-003, UR-TDS-M1-002, UR-TDS-M1-002, UR-TUA-M1-003, UR-DAC-M1-004, UR-DAC-M1-007, UR-DAC-M1-008, UR-TUA-PA-001, UR-TUA-PA-002 , UR-TUA-PA-003, UR-TUA-PA-004, UR-TDS-PA-010, UR-TUA-RT-001, UR-TDS-RT-012		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	191 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Rationale	It will enforce authenticity- the owner uses the private key to generate a digital signature, which will serve as a unique identifier of the sender and verify the integrity of the message.
-----------	--

Table 102: Self-sovereign identity management: FR-TRUST-SSIM-003

ID	FR-TRUST-SSIM-003	Author	UMU
Name	Signature verification		
Description	The TANGO dp-ABC cryptographic system must be able to verify signatures with the public key.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TDS-HT-010, UR-DAC-AV-007, UR-TUA-AV-001, UR-DAC-M1-008, UR-DAC-M1-007, UR-DAC-M1-004, UR-TUA-M1-003, UR-TDS-M1-002, UR-TDS-M1-002, UR-TUA-M1-003, UR-DAC-M1-004, UR-DAC-M1-007, UR-DAC-M1-008, UR-TUA-PA-001, UR-TUA-PA-002 , UR-TUA-PA-003, UR-TUA-PA-004 , UR-TDS-PA-010, UR-TUA-RT-001, UR-TDS-RT-012		
Rationale	Corresponding public key will be used to verify the digital signature so that it is ensured that the message was indeed sent by the claimed sender and there has been no tampering during the data sharing.		

Table 103: Self-sovereign identity management: FR-TRUST-SSIM-004

ID	FR-TRUST-SSIM-004	Author	UMU
Name	Token generation		
Description	The TANGO dp-ABC cryptographic system must generate a token corresponding to the selective disclosure ZK proof.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TDS-HT-010, UR-DAC-AV-007, UR-TUA-AV-001, UR-DAC-M1-008, UR-DAC-M1-007, UR-DAC-M1-004, UR-TUA-M1-003, UR-TDS-M1-002, UR-TDS-M1-002, UR-TUA-M1-003, UR-DAC-M1-004, UR-DAC-M1-007, UR-DAC-M1-008, UR-TUA-PA-001, UR-TUA-PA-002 , UR-TUA-PA-003, UR-TUA-PA-004 , UR-TDS-PA-010, UR-TUA-RT-001, UR-TDS-RT-012		
Rationale	Token generation serves as a cryptographic representation of the proof, containing only certain parts of the proof, while ensuring the confidentiality and integrity of the rest.		

Table 104: Self-sovereign identity management: FR-TRUST-SSIM-005

ID	FR-TRUST-SSIM-005	Author	UMU
Name	Token verification		
Description	The TANGO dp-ABC cryptographic system must be able to verify the Zkey token generated.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	192 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Priority	HIGH
Technology Category	TRUST
Related User Requirement	UR-TDS-HT-010, UR-DAC-AV-007, UR-TUA-AV-001, UR-DAC-M1-008, UR-DAC-M1-007, UR-DAC-M1-004, UR-TUA-M1-003, UR-TDS-M1-002, UR-TDS-M1-002, UR-TUA-M1-003, UR-DAC-M1-004, UR-DAC-M1-007, UR-DAC-M1-008, UR-TUA-PA-001, UR-TUA-PA-002, UR-TUA-PA-003, UR-TUA-PA-004, UR-TDS-PA-010, UR-TUA-RT-001, UR-TDS-RT-012
Rationale	Relevant information extracted from the token must be checked to ensure correctness and consistency of the proof by using public known parameters.

Table 105: Self-sovereign identity management: FR-TRUST-SSIM-006

ID	FR-TRUST-SSIM-006	Author	VTT
Name	Granular disclosure of personal data		
Description	The TANGO system should give users granular control over the amount of personal information revealed about them, for instance by considering what information is revealed by combining multiple data disclosures made at different times.		
Priority	MEDIUM		
Technology Category	TRUST		
Related User Requirement	UR-DMN-HT-002, UR-DMN-HT-004, UR-DMN-HT-005, UR-DMN-HT-008, UR-DAR-HT-009, UR-TDS-HT-010, UR-PCY-HT-012, UR-DMN-HT-013, UR-TDS-HT-019, UR-DAC-AV-007, UR-TDS-AV-008, UR-PCY-AV-009, UR-DAC-AV-010, UR-TDS-AV-011, UR-PCY-AV-012, UR-DMN-AV-013, UR-DMN-AV-014, UR-TDS-M1-002, UR-DAC-M1-004, UR-DST-M1-005, UR-DAC-M1-007, UR-DAC-M1-008, UR-DAR-M1-009, UR-GCO-M1-013, UR-DAR-M1-019, UR-PCY-M1-014, UR-DMN-M1-017, UR-DAR-M1-019, UR-GCO-M2-007, UR-DMN-M2-01, UR-DMN-RT-002, UR-TDS-RT-012, UR-GCO-RT-017, UR-PCY-RT-018, UR-DMN-RT-022, UR-TDS-PA-010, UR-GCO-PA-011, UR-PCY-PA-012, UR-PCY-PA-013		
Rationale	This feature will guarantee that no unnecessary user data is revealed, which is necessary according to the data security and privacy principles, and also simplifies GDPR compliance.		

Table 106: Self-sovereign identity management: FR-TRUST-SSIM-007

ID	FR-TRUST-SSIM-007	Author	
Name	Wallet (Personal Credential Manager)		
Description	The TANGO system must provide a wallet or a manager for personal credentials. This is an independent part of the whole SSI module.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-DMN-HT-002, UR-DMN-HT-004, UR-DMN-HT-005, UR-DMN-HT-008, UR-DAR-HT-009, UR-TDS-HT-010, UR-PCY-HT-012, UR-DMN-HT-013, UR-		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment		Page:	193 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	TDS-HT-019, UR-DAC-AV-007, UR-TDS-AV-008, UR-PCY-AV-009, UR-DAC-AV-010, UR-TDS-AV-011, UR-PCY-AV-012, UR-DMN-AV-013, UR-DMN-AV-014, UR-TDS-M1-002, UR-DAC-M1-004, UR-DST-M1-005, UR-DAC-M1-007, UR-DAC-M1-008, UR-DAR-M1-009, UR-GCO-M1-013, UR-DAR-M1-019, UR-PCY-M1-014, UR-DMN-M1-017, UR-DAR-M1-019, UR-GCO-M2-007, UR-DMN-M2-01, UR-DMN-RT-002, UR-TDS-RT-012, UR-GCO-RT-017, UR-PCY-RT-018, UR-DMN-RT-022, UR-TDS-PA-010, UR-GCO-PA-011, UR-PCY-PA-012, UR-PCY-PA-013
Rationale	This feature will provide means to store, delete and otherwise manage Personal Credentials for the use by SSI.

Table 107: Self-sovereign identity management: FR-TRUST-SSIM-008

ID	FR-TRUST-SSIM-008	Author	VTT
Name	Mechanism for detecting data misuse		
Description	The TANGO system should provide methods for detecting any use of data contrary to what was agreed to, for instance storing the data for longer time than agreed.		
Priority	MEDIUM		
Technology Category	TRUST		
Related User Requirement	UR-DMN-HT-002, UR-DMN-HT-004, UR-DMN-HT-005, UR-DMN-HT-008, UR-DAR-HT-009, UR-TDS-HT-010, UR-PCY-HT-012, UR-DMN-HT-013, UR-TDS-HT-019, UR-DAC-AV-007, UR-TDS-AV-008, UR-PCY-AV-009, UR-DAC-AV-010, UR-TDS-AV-011, UR-PCY-AV-012, UR-DMN-AV-013, UR-DMN-AV-014, UR-TDS-M1-002, UR-DAC-M1-004, UR-DST-M1-005, UR-DAC-M1-007, UR-DAC-M1-008, UR-DAR-M1-009, UR-GCO-M1-013, UR-DAR-M1-019, UR-PCY-M1-014, UR-DMN-M1-017, UR-DAR-M1-019, UR-GCO-M2-007, UR-DMN-M2-01, UR-DMN-RT-002, UR-TDS-RT-012, UR-GCO-RT-017, UR-PCY-RT-018, UR-DMN-RT-022, UR-TDS-PA-010, UR-GCO-PA-011, UR-PCY-PA-012, UR-PCY-PA-013		
Rationale	This feature will help increase user trust in how data is handled, which is necessary according to the data security and privacy principles.		

Table 108: Self-sovereign identity management: FR-TRUST-SSIM-009

ID	FR-TRUST-SSIM-009	Author	VTT
Name	Repudiability of user data disclosures		
Description	The TANGO system could provide methods for repudiation of disclosed data, to minimize the risk of user data being used in harmful ways.		
Priority	LOW		
Technology Category	TRUST		
Related User Requirement	UR-DMN-HT-002, UR-DMN-HT-004, UR-DMN-HT-005, UR-DMN-HT-008, UR-DAR-HT-009, UR-TDS-HT-010, UR-PCY-HT-012, UR-DMN-HT-013, UR-TDS-HT-019, UR-DAC-AV-007, UR-TDS-AV-008, UR-PCY-AV-009, UR-DAC-AV-010, UR-TDS-AV-011, UR-PCY-AV-012, UR-DMN-AV-013, UR-DMN-AV-		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	194 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	014, UR-TDS-M1-002, UR-DAC-M1-004, UR-DST-M1-005, UR-DAC-M1-007, UR-DAC-M1-008, UR-DAR-M1-009, UR-GCO-M1-013, UR-DAR-M1-019, UR-PCY-M1-014, UR-DMN-M1-017, UR-DAR-M1-019, UR-GCO-M2-007, UR-DMN-M2-01, UR-DMN-RT-002, UR-TDS-RT-012, UR-GCO-RT-017, UR-PCY-RT-018, UR-DMN-RT-022, UR-TDS-PA-010, UR-GCO-PA-011, UR-PCY-PA-012, UR-PCY-PA-013
Rationale	This feature will help minimize risks of data misuse, which is necessary according to the data security and privacy principles, and to increase user trust.

Table 109: Self-sovereign identity management: FR-TRUST-SSIM-010

ID	FR-TRUST-SSIM-010	Author	ATOS
Name	<i>Issue User with eID (simulating national eID)</i>		
Description	The TANGO system must issue eID Verifiable Credentials to end users to simulate the EU Digital Identity Wallet.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-HT-001, UR-TUA-AV-001,		
Rationale	This feature will guarantee that all users will be authenticated by their SSI Wallet credentials.		

Table 110: Self-sovereign identity management: FR-TRUST-SSIM-011

ID	FR-TRUST-SSIM-011	Author	ATOS
Name	<i>Issue User with Organization Member credential</i>		
Description	The TANGO system must issue Organization Members with Verifiable Credential to be able to act on behalf of their organization as according to their rôle/ job title.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-M1-001, UR-TUA-M1-003, UR-TUA-M2-001, UR-TUA-M2-002, UR-TUA-M2-005, UR-TUA-PA-001, UR-TUA-PA-002, UR-TUA-PA-003, UR-TUA-PA-004, UR-TUA-RT-001, UR-DUP-PA-007		
Rationale	This feature will guarantee that all users will be authenticated by their SSI Wallet credentials.		

Table 111: Self-sovereign identity management: FR-TRUST-SSIM-012

ID	FR-TRUST-SSIM-012	Author	ATOS
Name	<i>Verify Verifiable Credentials submitted to the Organizations</i>		
Description	The TANGO system must verify both eID and Organization Member Verifiable Credentials and return the identity claims in an ID Token.		
Priority	HIGH		
Technology Category	TRUST		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	195 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Related User Requirement	UR-TUA-M1-001, UR-TUA-M1-003, UR-TUA-M2-001, UR-TUA-M2-002, UR-TUA-M2-005, UR-TUA-PA-001, UR-TUA-PA-002, UR-TUA-PA-003, UR-TUA-PA-004, UR-TUA-RT-001
Rationale	This feature will guarantee that all users will be authenticated by their SSI Wallet credentials.

Table 112: Seamless Onboarding for Users and Devices: FR-TRUST-SOUD-001

ID	FR-TRUST-SOUD-001	Author	QBE
Name	Guest identity verification - Compatibility		
Description	The TANGO system must allow hotel guests to perform remote identity verification.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-DMN-HT-002		
Rationale	This feature will guarantee that the guests are the ones they claim to access the hotel, which is necessary according to the data security and privacy principles.		

Table 113: Seamless Onboarding for Users and Devices: FR-TRUST-SOUD-002

ID	FR-TRUST-SOUD-002	Author	QBE
Name	Applicants' identity verification		
Description	The TANGO system must allow visa applicants to perform remote identity verification.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-PA-001		
Rationale	This feature will guarantee that the identity of the applicants is correct, which is necessary according to the data security and privacy principles.		

Table 114: User continuous behavioural authentication: FR-TRUST-UCBA-001

ID	FR-TRUST-UCBA-001	Author	QBE
Name	User continuous behavioural authentication for hotel guests		
Description	The TANGO system must guarantee that the hotel guest accessing the platform, is who they claim they are.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-HT-001		
Rationale	This feature will guarantee that only authorised users will access the platform, which is necessary according to the data security and privacy principles.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	196 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 115: User continuous behavioural authentication: FR-TRUST-UCBA-002

ID	FR-TRUST-UCBA-002	Author	QBE
Name	User continuous behavioural authentication for Manufacturing		
Description	The TANGO system must guarantee that the person (staff employee) accessing the platform, is who they claim they are.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-M1-001, UR-TUA-M1-003		
Rationale	This feature will guarantee that only authorised users will access the platform, which is necessary according to the data security and privacy principles.		

Table 116: User continuous behavioural authentication: FR-TRUST-UCBA-003

ID	FR-TRUST-UCBA-003	Author	QBE
Name	User continuous behavioural authentication for Public Administration		
Description	The TANGO system must guarantee that the person (visa application and organisation employee) accessing the platform, is who they claim they are.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-PA-001, UR-TUA-PA-002, UR-TUA-PA-003, UR-TUA-PA-004		
Rationale	This feature will guarantee that only authorised users will access remotely, which is necessary according to the data security and privacy principles.		

Table 117: User continuous behavioural authentication: FR-TRUST-UCBA-004

ID	FR-TRUST-UCBA-004	Author	QBE
Name	User continuous behavioural authentication for Retail		
Description	The TANGO system must guarantee that the person accessing the platform, is who they claim they are.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-RT-001		
Rationale	This feature will guarantee that only authorised users will access remotely, which is necessary according to the data security and privacy principles.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	197 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 118: User continuous behavioural authentication: FR-TRUST-UCBA-005

ID	FR-TRUST-UCBA-005	Author	QBE
Name	Transaction fraud detection		
Description	The TANGO system must be able to detect fraud in different transactions.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-PMT-BN-002, UR-SAD-BN-003		
Rationale	This feature will allow the detection of fraud based on different types of transactions.		

Table 119: Device continuous behavioural authentication: FR-TRUST-DCBA-001

ID	FR-TRUST-DCBA-001	Author	UTH
Name	Device hardware information monitoring		
Description	The TANGO system must be able to monitor, poll and log the device's hardware information (id, manufacturer, model, processor type, speed, etc.).		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002		
Rationale	This data will assist in determining whether the device is similar to the one that is authenticated.		

Table 120: Device continuous behavioural authentication: FR-TRUST-DCBA-002

ID	FR-TRUST-DCBA-002	Author	UTH
Name	Device network status and network traffic monitoring		
Description	The TANGO system must be able to monitor, poll and log the device's network status (IMEI number, IP address, MAC address, WLAN SSID, etc.) as well as the device's network traffic patterns (DNS queries, packet size and timing, port usage etc.).		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002		
Rationale	This data will assist in determining whether the device exhibits similar behavioral patterns to the one that is authenticated.		

Table 121: Device continuous behavioural authentication: FR-TRUST-DCBA-003

ID	FR-TRUST-DCBA-003	Author	UTH
Name	Device location familiarity monitoring		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	198 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Description	The TANGO system must be able to monitor, poll and log the device's location familiarity (nearby RF signals such as WiFi, Bluetooth devices, etc.).
Priority	HIGH
Technology Category	TRUST
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002
Rationale	This data will assist in determining whether the location of the authenticated device has been changed.

Table 122: Device continuous behavioural authentication: FR-TRUST-DCBA-004

ID	FR-TRUST-DCBA-004	Author	UTH
Name	Data preprocessing		
Description	The TANGO system should be able to render extracted data in a predefined and standard format, handle invalid entries and undefined values and prepare the processes data for feeding into AI/ML models for behavioural analysis and classification.		
Priority	MEDIUM		
Technology Category	TRUST		
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002		
Rationale	The preprocessing of data into a pre-defined and standard format will enable the deployment of DCBA onto multiple, heterogeneous devices and allow the training of the AI/ML models with consistent and homogenized data.		

Table 123: Device continuous behavioural authentication: FR-TRUST-DCBA-005

ID	FR-TRUST-DCBA-005	Author	UTH
Name	AI/ML model for behavioural analysis		
Description	The TANGO system must be able to feed the preprocessed data to an AI/ML model for behavioral recognition and anomaly detection.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002		
Rationale	An AI/ML model trained on the device's data will be able to establish a baseline of device "normal" behavior and further classify any subsequent data from the same device as "normal" or "abnormal". This functionality will assist in determining if a new device authentication process must occur.		

Table 124: Device continuous behavioural authentication: FR-TRUST-DCBA-006

ID	FR-TRUST-DCBA-006	Author	UTH
----	-------------------	--------	-----

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	199 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Name	Alert notification
Description	The TANGO system must be able to generate alerts based on the verdict of the AI model for abnormal behavior, and provide contextual information, including the affected device along with details of the behavior change that triggered the alert.
Priority	HIGH
Technology Category	TRUST
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002, UR-PMT-M2-003, UR-PMT-M2-004, UR-PMT-M2-008, UR-DAR-M2-009, UR-PMT-M2-010, UR-TUA-M2-012, UR-TUA-M2-013
Rationale	An alert mechanism will allow the triggering of a new SSI authentication process, as well as provide awareness and information to administrators and supervisors of abnormal behavior.

Table 125: Hardening against side-channel attacks: FR-TRUST-HASA-001

ID	FR-TRUST-HASA-001	Author	CEA
Name	Side-channel attack resilience		
Description	The smart hospitality components for sensor data ciphering physically accessible to an attacker must include side-channel countermeasures.		
Priority	HIGH		
Technology Category	TRUST		
Related User Requirement	UR-DMN-HT-004, UR-PCY-HT-012		
Rationale	Our compiler will apply countermeasures against side-channel attacks to the crypto implementation used for the communication between the sensors in the rooms with the cloud platform. This feature will avoid encryption keys used for the communication to be found via side-channel attacks. Such attacks would break confidentiality, authenticity and integrity of data.		

B.2 Non-Functional Requirements

Table 126: Self-sovereign identity management: NFR-SEC-SSIM-001

ID	NFR-SEC-SSIM-001	Author	UMU
Name	Distributed issuance		
Description	The TANGO dp-ABC cryptographic system could enable distributed issuance processes, through the aggregation of public keys and combination of signatures into a single one.		
Priority	LOW		
Technology Category	SECURITY		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	200 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Related User Requirement	UR-TDS-HT-010, UR-DAC-AV-007, UR-TUA-AV-001, UR-DAC-M1-008, UR-DAC-M1-007, UR-DAC-M1-004, UR-TUA-M1-003, UR-TDS-M1-002, UR-TDS-M1-002, UR-TUA-M1-003, UR-DAC-M1-004, UR-DAC-M1-007, UR-DAC-M1-008, UR-TUA-PA-001, UR-TUA-PA-002, UR-TUA-PA-003, UR-TUA-PA-004, UR-TDS-PA-010, UR-TUA-RT-001, UR-TDS-RT-012
Rationale	Distributed issuance is a powerful tool that might be useful in some scenarios, to enable further security guarantees. However, it is not necessary for the use of zero-knowledge technologies for improvement of privacy and security in the authentication process.

Table 127: Self-sovereign identity management: NFR-SEC-SSIM-002

ID	NFR-SEC-SSIM-002	Author	UMU
Name	Confidentiality insurance		
Description	The TANGO dp-ABC cryptographic system must ensure the confidentiality of sensitive information.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-TUA-HT-001, UR-TDS-HT-010, UR-TUA-AV-001, UR-PCY-AV-003, UR-PCY-AV-004, UR-TUA-M1-001, UR-TDS-M1-002, UR-TUA-M2-001, UR-TUA-M2-002, UR-TUA-M2-005, UR-TUA-PA-001, UR-TUA-PA-002, UR-TUA-PA-003, UR-TUA-PA-004, UR-TUA-RT-001, UR-PCY-RT-018, UR-PCY-RT-019		
Rationale	The system must maintain the confidentiality of sensitive information, I.e., private keys, proof details. It must prevent unauthorized access, disclosure or leakage of information.		

Table 128: Self-sovereign identity management: NFR-SEC-SSIM-003

ID	NFR-SEC-SSIM-003	Author	UMU
Name	Integrity insurance		
Description	The TANGO dp-ABC cryptographic system must ensure the integrity of sensitive information.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-TDS-HT-010, UR-DIN-AV-002, UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-AV-009, UR-TDS-AV-011, UR-TDS-M1-002, UR-PCY-M1-014, UR-DUP-PA-005, UR-PCY-PA-013, UR-PCY-RT-018, UR-PCY-RT-019		
Rationale	The system must detect any tampering attempts on the proof.		

Table 129: Self-sovereign identity management: NFR-SEC-SSIM-004

ID	NFR-SEC-SSIM-004	Author	UMU
Name	Authenticity insurance		
Description	The TANGO dp-ABC cryptographic system must ensure the authenticity of proofs.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment		Page:	201 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Priority	HIGH
NFR Category	SECURITY
Related User Requirement	UR-TUA-HT-001, UR-TDS-HT-010, UR-TUA-AV-001, UR-PCY-AV-003, UR-PCY-AV-004, UR-TUA-M1-001, UR-TDS-M1-002, UR-TUA-M1-003, UR-DAC-M1-008, UR-TUA-M2-002, UR-PCY-PA-013, UR-PCY-RT-018, UR-PCY-RT-019
Rationale	The system must provide authentication mechanisms to ensure that proof originates from legitimate sources.

Table 130: Self-sovereign identity management: NFR-PER-SSIM-001

ID	NFR-PER-SSIM-001	Author	UMU
Name	Scalability and responsiveness		
Description	The TANGO dp-ABC cryptographic system should scale effectively as the number of proof requests and system participants increases.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		
Related User Requirement	UR-DUP-AV-005, UR-DAC-AV-010, UR-DMN-RT-002		
Rationale	The system should handle a growing workload without a significant decrease in performance.		

Table 131: Self-sovereign identity management: NFR-REL-SSIM-001

ID	NFR-REL-SSIM-001	Author	UMU
Name	Accuracy and Correction		
Description	The TANGO dp-ABC cryptographic system must produce accurate and correct outputs.		
Priority	HIGH		
NFR Category	RELIABILITY		
Related User Requirement	UR-TDS-HT-010, UR-DIN-AV-002, UR-PCY-AV-003, UR-PCY-AV-004, UR-DUP-PA-005, UR-PCY-PA-013		
Rationale	The system must output accurate and correct results during proof generation and verification. Proofs must represent accurately and correctly the statements to be proved.		

Table 132: Self-sovereign identity management: NFR -REL-SSIM-002

ID	NFR-REL-SSIM-002	Author	UMU
Name	Fault Tolerance and Availability		
Priority	MEDIUM		
Description	The system should be highly available, allowing users to access the method without significant downtime or interruptions in case of faults or failures.		
NFR Category	RELIABILITY		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	202 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Related User Requirement	UR-DUP-AV-005, UR-DAC-AV-010, UR-DAC-M1-004, UR-DMN-RT-002, UR-PCY-RT-018, UR-PCY-RT-019
--------------------------	---

Table 133: Self-sovereign identity management: NFR-COMP-SSIM-001

ID	NFR-COMP-SSIM-001	Author	ATOS
Name	SSI Agent to follow the Self-Issued OpenID Provider v2 standard		
Description	The SSI Agent must follow the GDPR to respect the privacy of the user's personal data.		
Priority	HIGH		
NFR Category	COMPATIBILITY		
Rationale	The reason for implementing SIOP is to follow the Grant Agreement requirement and also to align with EBSI and the EU Digital Identity Wallet. Furthermore, this aligns with Data Spaces Technical Convergence.		

Table 134: Self-sovereign identity management: NFR-SEC-SSIM-005

ID	NFR-SEC-SSIM-005	Author	ATOS
Name	SSI to follow the GDPR		
Description	The SSI Agent must comply with the GDPR and any related regulations as identified by the pilot partners.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-GCO-HT-011, UR-GCO-AV-006, UR-GCO-M1-013, UR-GCO-M2-007, UR-GCO-PA-011, UR-GCO-RT-017		
Rationale	The reason for complying with the GDPR and any related regulations is to satisfy legal compliance and provide a privacy preserving solution to the end user.		

Table 135: Seamless Onboarding for Users and Devices: NFR-COMP-SOUD-001

ID	NFR-COMP-SOUD-001	Author	QBE
Name	Guest identity verification - Compatibility		
Description	The TANGO system must be compatible with SSI		
Priority	HIGH		
NFR Category	COMPATIBILITY		
Related User Requirement	UR-DMN-HT-002, UR-DMN-HT-005, UR-TUA-PA-001		
Rationale	Guests must be able to securely perform remote identity verification before they arrive to the hotel.		

Table 136: Seamless Onboarding for Users and Devices: NFR-PER-SOUD-001

ID	NFR-PER-SOUD-001	Author	QBE
Name	Guest identity verification – Performance/Reliability/Security		
Description	The TANGO system must support identity verification in a performant, reliable, secure manner		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	203 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Priority	HIGH
NFR Category	PERFORMANCE / RELIABILITY / SECURITY
Related User Requirement	UR-DMN-HT-002
Rationale	Guests have to be able to securely perform remote identity verification before they arrive to the hotel. The process has to take place in a timely manner without creating frustration to the user and also be available when the user is trying to access the service

Table 137: Seamless Onboarding for Users and Devices: NFR-EXIST-SOUD-001

ID	NFR-EXIST-SOUD-001	Author	QBE
Name	Guest identity verification – HW Constraints		
Description	The TANGO system must perform the identity verification through smartphones.		
Priority	HIGH		
NFR Category	HW & SW CONSTRAINTS		
Related User Requirement	UR-DMN-HT-002, UR-DMN-HT-005, UR-TUA-PA-001		
Rationale	Guests have to be able to securely perform remote identity verification before they arrive to the hotel.		

Table 138: User continuous behavioural authentication: NFR-COMP-UCBA-001

ID	NFR-COMP-UCBA-001	Author	QBE
Name	Compatibility of User continuous behavioural authentication		
Description	The solution must be compatible with the TANGO mobile and web apps.		
Priority	HIGH		
NFR Category	COMPATIBILITY		
Rationale	Allow integration with the mobile apps of the TANGO platform both for mobile and web.		

Table 139: User continuous behavioural authentication: NFR-REL-UCBA-001

ID	NFR-REL-UCBA-001	Author	QBE
Name	Real-time and reliability constraints of User continuous behavioural authentication		
Description	The TANGO system must ensure that the user continuous behavioural authentication will operate in real-time, reliably and to be secure.		
Priority	HIGH		
NFR Category	PERFORMANCE / RELIABILITY / SECURITY		
Rationale	The response time should be at most 2 seconds, to be robust in real environments, secure against known adversarial attacks.		

Table 140: User continuous behavioural authentication: NFR-EXIST-UCBA-001

ID	NFR-EXIST-UCBA-001	Author	QBE
----	--------------------	--------	-----

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	204 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Name	User continuous behavioural authentication is provided for smartphones and web apps
Description	The TANGO system must support smartphone and web apps.
Priority	HIGH
NFR Category	HW & SW CONSTRAINTS
Rationale	The solution will need to support smartphones and web apps, considering that the users will interact with the TANGO platform in the different use cases through different interfaces, in some cases mobile in some cases web apps.

Table 141: User continuous behavioural authentication: NFR-REL-UCBA-002

ID	NFR-REL-UCBA-002	Author	QBE
Name	Performance and Reliability of the transaction behaviour detection mechanism.		
Description	The TANGO system should be able to accurately and reliably detect fraud based on different types of transactions.		
Priority	MEDIUM		
NFR Category	PERFORMANCE / RELIABILITY		
Related User Requirement	UR-PMT-007, UR-PMT-008		
Rationale	The component should be able to accurately and reliable detect fraud based on different types of transactions.		

Table 142: Device continuous behavioural authentication: NFR-COMP-DCBA-001

ID	NFR-COMP-DCBA-001	Author	UTH
Name	DBCA compatibility		
Description	The TANGO system should ensure compatibility and applicability of the DBCA module.		
Priority	MEDIUM		
NFR Category	COMPATIBILITY / PORTABILITY		
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002		
Rationale	The DBCA module should be compatible with a wide range of devices, operating systems, and hardware configurations. It should also integrate with popular data streaming software and protocols. All software components should be containerized in order to enhance portability.		

Table 143: Device continuous behavioural authentication: NFR-PER-DCBA-001

ID	NFR-PER-DCBA-001	Author	UTH
Name	Effectiveness and responsiveness		
Description	The TANGO system should ensure the efficient execution of the DBCA module, as well as the rapid reaction to abnormal behavior.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	205 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002
Rationale	The data extraction agents should efficiently monitor, poll, log and stream device information without significant impact on device performance and the AI/ML model for behavioral analysis should be optimized for fast and accurate predictions to support real-time anomaly detection.

Table 144: Device continuous behavioural authentication: NFR-REL-DCBA-001

ID	NFR-REL-DCBA-001	Author	UTH
Name	Accuracy and fault-tolerance		
Description	The TANGO system must ensure continuous and accurate device behavioral authentication.		
Priority	HIGH		
NFR Category	RELIABILITY		
Related User Requirement	UR-TUA-AV-001, UR-DUP-AV-005, UR-DAC-AV-007, UR-DAC-AV-010, UR-DMN-AV-013, UR-DMN-AV-014, UR-TUA-M2-001, UR-TUA-M2-002		
Rationale	The module must be highly reliable, ensuring continuous monitoring and extraction of device data, as well as handle unexpected errors, exceptions and network disruptions without compromising the overall functionality. Furthermore, the AI/ML model should exhibit high accuracy and robustness in detecting behavioral anomalies.		

Table 145: Device continuous behavioural authentication: NFR-SEC-DCBA-001

ID	NFR-SEC-DCBA-001	Author	UTH
Name	Privacy and GDPR compliance		
Description	The TANGO system must ensure that personal data is not leaked, and data handling is GDPR-compliant.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-DUP-AV-005, UR-GCO-AV-006, UR-PCY-M2-006, UR-GCO-M2-007		
Rationale	Every developed software component must implement robust security measures to protect sensitive information and user privacy (i.e., anonymization), while secure communication protocols (e.g., HTTPS) must be employed for data transmission between components. Finally, access controls and encryption techniques must be implemented to safeguard the integrity and confidentiality of data.		

Table 146: Hardening against side-channel attacks: NFR-EXIST-HASA-001

ID	NFR-EXIST-HASA-001	Author	CEA
Name	Compatibility between compiler that applies countermeasures and source code and target platform.		
Description	The compiler developed to automate the application of countermeasures against side-channel attacks requires C code as source code, and requires the target platform		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	206 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	to be an ARMv7M platform. Thus, source code for encryption must be written in C, and the platform used must have an ARMv7M-compatible CPU.
Priority	HIGH
NFR Category	COMPATIBILITY / HW & SW CONSTRAINTS
Related User Requirement	UR-DMN-HT-004, UR-PCY-HT-012
Rationale	This requirement is due to constraints from existing code base. The compiler already supports code polymorphism for ARMv7M and C code, and it will be extended during the project.

Table 147: Hardening against side-channel attacks: NFR-COMP-HASA-001

ID	NFR-COMP-HASA-001	Author	CEA
Name	Code polymorphism and memory permissions.		
Description	The code polymorphism countermeasure relies on dynamic code generation: write and execute permissions should be given to a memory segment in order to deploy this countermeasure.		
Priority	MEDIUM		
NFR Category	COMPATIBILITY / HW & SW CONSTRAINTS		
Related User Requirement	UR-DMN-HT-004, UR-PCY-HT-012		
Rationale	Without memory section both writable and executable, the code polymorphism countermeasure cannot be used. It consists in calling regularly at runtime a code generator to make the code of a function vary. So, the generator must be able to write the code in memory, and we must be able to execute it afterwards.		

Table 148: Hardening against side-channel attacks: NFR-SEC-HASA-001

ID	NFR-SEC-HASA-001	Author	CEA
Name	Reasonable performance/security tradeoff.		
Description	Securing against side-channel attacks forces to find a trade-off between performance (i.e. clock cycles) and security (i.e. how resistant to an attack the code will be). Security must be as good as possible for a given performance constraint.		
Priority	HIGH		
NFR Category	PERFORMANCE / SECURITY		
Related User Requirement	UR-DMN-HT-004, UR-PCY-HT-012		
Rationale	Reaching perfect security against side-channel attack is not always possible, especially in software. Thus, a trade-off has to be found, depending on both security and performance requirements.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	207 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

C. System Requirements for AI-based Framework for Green & Trustworthy Operations

C.1 Functional Requirements

Table 149: Exploratory Data Analysis Engine: FR-AI-EDAE-001

ID	FR-AI-EDAE-001	Author	EXUS
Name	EDAE_customer_categorisation		
Description	The TANGO system must divide customers into categories based on shopping frequency, amounts spent and recency and generate relevant reports		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-RT-007, UR-DAR-RT-014, UR-DAR-RT-016		
Rationale	The purpose of this FR is to define customers' loyalty levels		

Table 150: Exploratory Data Analysis Engine: FR-AI-EDAE-002

ID	FR-AI-EDAE-002	Author	EXUS
Name	EDAE_pattern_recognition		
Description	The TANGO system must detect and report shopping patterns and preferences between customers of different countries using different systems and structures.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-RT-007, UR-DAR-RT-014, UR-DAR-RT-016		
Rationale	The purpose of this requirement is to support the end user to identify patterns and preferences among different countries and to integrate relevant data into one platform.		

Table 151: Exploratory Data Analysis Engine: FR-AI-EDAE-003

ID	FR-AI-EDAE-003	Author	EXUS
Name	EDAE_new_customer_prospects		
Description	The TANGO system must expose data to analysts of suppliers, new customer prospects that are setting up their business case, other markets, and the tourism industry.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-RT-007, UR-DAR-RT-014, UR-DAR-RT-016		
Rationale	The purpose of this requirement is to identify and report new market opportunities that occur.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	208 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 152: Exploratory Data Analysis Engine: FR-AI-EDAE-004

ID	FR-AI-EDAE-004	Author	EXUS
Name	EDAE_deep_analysis		
Description	The TANGO system must provide a deep analysis of the sales data, product data, and customer data.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-RT-007, UR-DAR-RT-014, UR-DAR-RT-016		
Rationale	Self-explanatory.		

Table 153: Exploratory Data Analysis Engine: FR-AI-EDAE-005

ID	FR-AI-EDAE-005	Author	EXUS
Name	EDAE_weak_areas		
Description	The TANGO system must spot weak areas in a store.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-RT-007, UR-DAR-RT-014, UR-DAR-RT-016		
Rationale	The purpose of this requirement is to identify areas that can be targeted for increased revenue.		

Table 154: Exploratory Data Analysis Engine: FR-AI-EDAE-006

ID	FR-AI-EDAE-006	Author	EXUS
Name	EDAE_customer_behaviour		
Description	The TANGO system should explore large datasets pertaining to customer purchasing patterns.		
Priority	MEDIUM		
Technology Category	AI		
Related User Requirement	UR-DMN-RT-007, UR-DAR-RT-014, UR-DAR-RT-016		
Rationale	The purpose of this requirement is to indicate regional taste preferences and provide a range of insights into customer behaviour.		

Table 155: Exploratory Data Analysis Engine: FR-AI-EDAE-007

ID	FR-AI-EDAE-007	Author	EXUS
Name	EDAE_statistics_and_analysis_reports		
Description	The TANGO system must provide hotel (group) managers with statistics and analysis reports on the data collected across all hotels.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-HT-009		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	209 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Rationale	This requirement derives from a relevant request from the end user and will address a challenge they currently face.
-----------	--

Table 156: Exploratory Data Analysis Engine: FR-AI-EDAE-008

ID	FR-AI-EDAE-008	Author	EXUS
Name	EDAE_customer_trends_and_patterns		
Description	The TANGO system must inform hotel (group) managers on patterns, trends and associations between a client's traits and preferences and those of other clients.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-HT-009		
Rationale	This will help hotel managers to predict client's preferences based on past experiences.		

Table 157: Exploratory Data Analysis Engine: FR-AI-EDAE-009

ID	FR-AI-EDAE-009	Author	EXUS
Name	EDAE_print_reports		
Description	The TANGO system must provide technicians with reports on the print (incl. failed) jobs to improve the printer operation and the printing quality.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-010		
Rationale	This will assist technicians to gather information on failed parts to understand why they might have failed.		

Table 158: Exploratory Data Analysis Engine: FR-AI-EDAE-010

ID	FR-AI-EDAE-010	Author	EXUS
Name	EDAE_customer_data		
Description	The TANGO system must provide managers with customers' data to get KPIs from them.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-011		
Rationale	The gathering of customers' data will assist managers to set KPIs for the new print jobs.		

Table 159: Exploratory Data Analysis Engine: FR-AI-EDAE-011

ID	FR-AI-EDAE-011	Author	EXUS
Name	EDAE_quality_issues		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	210 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Description	The TANGO system should provide data scientists with data related to prints' quality issues.
Priority	MEDIUM
Technology Category	AI
Related User Requirement	UR-DAR-M1-018
Rationale	Self-explanatory.

Table 160: Exploratory Data Analysis Engine: FR-AI-EDAE-012

ID	FR-AI-EDAE-012	Author	EXUS
Name	EDAE_data_fusion		
Description	The TANGO system must be able to provide users with tools to perform fusion from multiple sources (possibly with different data structures/schemas) and visualize the results.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-RT-003, UR-DMN-RT-020		
Rationale	Data fusion is an essential process when dealing with multisource data to improve its quality and facilitate ML operations.		

Table 161: Exploratory Data Analysis Engine: FR-AI-EDAE-013

ID	FR-AI-EDAE-013	Author	EXUS
Name	EDEA_notifications		
Description	The TANGO system must be able to generate notifications regarding the status of all EDAE processing tasks that will be sent to the visualization and the UI components.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-RT-013		
Rationale	Notifying the users about the progress and after the completion of the analysis tasks improves the overall user experience.		

Table 162: Federated Learning Component: FR-AI-FL-001

ID	FR-AI-FL-001	Author	ATOS
Name	Federated learning training node setup		
Description	The TANGO system must provide the capability to setup a data processing machine as a training node of a federated learning architecture.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	211 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Rationale	TANGO must provide the functionality of setting up a data processing machine with a dataset stored locally as a federated learning node in order to train a given model with the available node in that dataset. The federated learning training node must have connectivity with the federated learning infrastructure main node.
-----------	--

Table 163: Federated Learning Component: FR-AI-FL-002

ID	FR-AI-FL-002	Author	ATOS
Name	Federated learning main node setup		
Description	The TANGO system must provide the capability to setup a data processing machine as a main node of a federated learning architecture.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	For each TANGO scenario a main node to aggregate the results of the different training outcomes coming from the different training nodes must be configured.		

Table 164: Federated Learning Component: FR-AI-FL-003

ID	FR-AI-FL-003	Author	ATOS
Name	Provide support for python programming language		
Description	The TANGO system should support the usage of python programming language in the federated learning process.		
Priority	MEDIUM		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	Provided Federated learning solution might provide support for the python programming language and some of its ML frameworks in order to perform the training of the models.		

Table 165: MLOps Component: FR-AI-MLOPS-001

ID	FR-AI-MLOPS-001	Author	ATOS
Name	Model tracking		
Description	The TANGO system must offer the functionality to keep track of different model versions generated during experimentations.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020, UR-SAD-BN-003		
Rationale	Those users of the TANGO platform that are interested on the generation and experimentation with Machine Learning predictive models will be offered the functionality of keeping track of the different models generated on their experimentations, storing also the metrics of		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	212 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	interests related to those models in order to get insights and comparison in order to evaluate the experimentation process.
--	---

Table 166: MLOps Component: FR-AI-MLOPS-002

ID	FR-AI-MLOPS-002	Author	ATOS
Name	Model deployment		
Description	The TANGO system should be able to allow the deployment of models that have been persisted in the MLOps component.		
Priority	MEDIUM		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	In order to use models tracked by the MLOps component, this component should allow the deployment of trained models into production environments.		

Table 167: AutoML Component: FR-AI-AUTOML-001

ID	FR-AI-AUTOML-001	Author	ATOS
Name	Automated model optimization		
Description	The TANGO system must provide a mechanism to generate a model resultant from the automatic optimization of its parameters.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020, UR-SAD-BN-003		
Rationale	For those use cases where the creation of a predictive model is needed, this component will allow the creation of a trained model with its hyperparameters optimized to generate better results for the metrics selected with a given training dataset.		

Table 168: AutoML Component: FR-AI-AUTOML-002

ID	FR-AI-AUTOML-002	Author	ATOS
Name	Model evaluation		
Description	The TANGO system must provide metrics that allows the end user to be aware of the evaluation carried out during the model training to decide the best model to choose.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020, UR-DAR-M1-018, UR-DAR-BN-008		
Rationale	AutoML component must provide the metrics of the different models generated in order to have a clear view of the whole process of parameter optimization. Additionally, the user might have the capability to define		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	213 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	which metrics to prioritize in the optimization process in order to drive the strategy to follow by the component in the model generation.
--	--

Table 169: TornadoVM: FR-AI-TVM-001

ID	FR-AI-TVM-001	Author	UOM
Name	Code for data analytics must offer parallelism		
Description	The data analytics that will be developed by the EXUS EDAE engine in Task T5.1 must not use any data dependency.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	This functional requirement is indirectly mapped to UR-DAR-M1-009 and UR-DAR-M1-010 .		
Rationale	This requirement must be satisfied in order to be sensible the utilization of hardware acceleration. Otherwise, the code that will be delegated on a GPU will result in low performance and the energy consumption can be higher than running on a CPU.		

Table 170: TornadoVM, FR-AI-TVM-002

ID	FR-AI-TVM-002	Author	UOM
Name	Large size of data to feed the parallel processing on hardware accelerators		
Description	The data sizes that will be given for the analytics (developed by the EXUS EDAE engine in Task T5.1) must be beyond the scale of 2k elements (array length).		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	This functional requirement is indirectly mapped to UR-DAR-M1-009 and UR-DAR-M1-010 .		
Rationale	This requirement must be satisfied in order to ensure that the cost of the data transfers via PCIe will not exceed any performance improvement performed by a hardware accelerator device (e.g., a GPU). For example, a GPU usually offers thousands of cores and the total amount of memory is in the scale of 4-16 GBs, depending on the platform model.		

Table 171: Privacy Enhancing Component: FR-AI-PEC-001

ID	FR-AI-PEC-001	Author	UoG
Name	PEC_privacy_assessment		
Description	The TANGO system must provide privacy impact assessment to identify critical data actions, threats and their possible privacy impact, and mitigation measures to be able to preserve privacy of users and assure trust on the system.		
Priority	HIGH		
Technology Category	AI		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	214 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Related User Requirement	UR-PCY-AV-003, UR-PCY-AV-004, UR-PCY-RT-018, UR-GCO-RT-019
Rationale	The purpose of this FR is to support in identifying sensitive data actions, privacy requirements, possible threats and their privacy impacts on users.

Table 172: Privacy Enhancing Component: FR-AI-PEC-002

ID	FR-AI-PEC-002	Author	UoG
Name	PEC_recommendations		
Description	The TANGO system must provide a list of recommendations to address the identified privacy risks.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-GCO-AV-006, UR-GCO-RT-017		
Rationale	The purpose of this FR is to identify possible countermeasures to mitigate privacy risks.		

Table 173: Privacy Enhancing Component: FR-AI-PEC-003

ID	FR-AI-PEC-003	Author	UoG
Name	PEC_compliance		
Description	The TANGO system could provide suggestions on complying with GDPR		
Priority	LOW		
Technology Category	AI		
Related User Requirement	UR-GCO-AV-006, UR-GCO-RT-017		
Rationale	The purpose of this FR is to identify possible countermeasures to assist with GDPR compliance.		

Table 174: Privacy Assurance Tool: FR-AI-PAT-001

ID	FR-AI-PAT-001	Author	FUJ_LU
Name	Collect client's personal data as well as their preferences		
Description	The TANGO system must be able to collect client's personal data directly through the web. In addition, it must be able to collect the clients' preferences to provide them with personalised services.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-HT-002, UR-DMN-HT-003		
Rationale	This feature will guarantee the collected client's personal data as well as their preferences directly through the web.		

Table 175: Privacy Assurance Tool: FR-AI-PAT-002

ID	FR-AI-PAT-002	Author	FUJ_LU
----	---------------	--------	--------

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	215 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Name	Trustworthy Data Sharing
Description	The TANGO system must enable data sharing only among the relevant stakeholders/business partners.
Priority	HIGH
Technology Category	AI
Related User Requirement	UR-TDS-HT-010
Rationale	Creation of a feature allowing data sharing only among the relevant stakeholders/business partners.

Table 176: Privacy Assurance Tool: FR-AI-PAT-003

ID	FR-AI-PAT-003	Author	FUJ_LU
Name	GDPR compliant-based system		
Description	The TANGO system must allow management of data based on GDPR requirements.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-GCO-HT-011		
Rationale	The system must be GDPR compliant.		

Table 177: Privacy Assurance Tool: FR-AI-PAT-004

ID	FR-AI-PAT-004	Author	FUJ_LU
Name	Authentication to access to the system		
Description	The TANGO system must provide a trusted user authentication for VISAR applicants.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-TUA-PA-001		
Rationale	This feature enables the support of trusted user authentication for applicants.		

Table 178: Privacy Assurance Tool: FR-AI-PAT-005

ID	FR-AI-PAT-005	Author	FUJ_LU
Name	Authentication– VISAR Employees		
Description	The TANGO system must provide a trusted user authentication for VISAR employees.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-TUA-PA-002		
Rationale	This feature enables the support of trusted user authentication for VISAR employees to access the system.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	216 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 179: Privacy Assurance Tool: FR-AI-PAT-006

ID	FR-AI-PAT-006	Author	FUJ_LU
Name	Trustworthy document upload		
Description	The TANGO system must provide a secure data uploading channel.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DUP-PA-005, UR-DUP-PA-006		
Rationale	This feature ensures safe document and data upload in user and organization side.		

Table 180: Privacy Assurance Tool: FR-AI-PAT-007

ID	FR-AI-PAT-007	Author	FUJ_LU
Name	Monitoring and managing the documents and the data flow in the system.		
Description	The TANGO system must provide a data management approach for monitoring and managing the shared data.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DMN-PA-009		
Rationale	This feature enables specific users to monitor and manage the data flow in the system.		

Table 181: XAI Component: FR-AI-XAI-001

ID	FR-AI-XAI-001	Author	SQD
Name	Explainability of Machine Learning models		
Description	The TANGO system must explain the model's predictions and describe which features had a great impact on creating the specific results.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-020, UR-DAR-BN-007		
Rationale	With the use of X-AI, the user must be able to understand the circumstances under which a ML model produced its results.		

Table 182: XAI Component: FR-AI-XAI-002

ID	FR-AI-XAI-002	Author	SQD
Name	Explainability of Machine Learning models for debugging		
Description	The TANGO system must explain the model's predictions, giving the data scientist the opportunity to easily identify problems during the model training.		
Priority	HIGH		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	217 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Technology Category	AI
Related User Requirement	UR-DAR-M1-020, UR-DAR-BN-008
Rationale	Using X-AI when training a model is highly recommended so the model can be quickly debugged since unexpected bias can easily be discovered.

Table 183: RENOPS Component: FR-AI-RENOPS-001

ID	FR-AI-RENOPS-001	Author	XLAB
Name	Prediction of renewable energy availability		
Description	The TANGO system must provide an accurate forecast of renewable energy availability.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	Prediction of availability of renewables will address the energy efficiency part of TANGO.		

Table 184: RENOPS Component: FR-AI-RENOPS-002

ID	FR-AI-RENOPS-002	Author	XLAB
Name	Accessibility of the model output		
Description	The TANGO RENOPS must provide the model prediction via an API.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	Making data accessible via API will enable us to use this information in related requirements and task		

Table 185: RENOPS Component: FR-AI-RENOPS-003

ID	FR-AI-RENOPS-003	Author	XLAB
Name	Task scheduling in time		
Description	The TANGO RENOPS system must be able to utilize information from the forecasting service and schedule energy intensive tasks in time.		
Priority	HIGH		
Technology Category	AI		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	Using the renewable availability forecast to find the optimal time to execute the energy intensive tasks will address the energy efficiency aspect of TANGO.		

Table 186: RENOPS Component: FR-AI-RENOPS-004

ID	FR-AI-RENOPS-004	Author	XLAB
Name	Task scheduling in space		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	218 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Description	The TANGO RENOPS system could utilize information from the forecasting service to geographically schedule the workload.
Priority	LOW
Technology Category	AI
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020
Rationale	Using the renewable availability forecast for multiple locations to find the optimal datacentre (HPC) to execute the energy tasks which will address the energy efficiency aspect of TANGO.

C.2 Non-Functional Requirements

Table 187: Exploratory Data Analysis Engine: NFR-COMP -EDAE-001

ID	NFR-COMP-EDAE-001	Author	EXUS
Name	Compatibility with other TANGO components		
Description	The TANGO system should be compatible with all TANGO components that consume its data and with any user platform that feeds the EDAE.		
Priority	MEDIUM		
NFR Category	COMPATIBILITY		
Rationale	Compatibility with other linked components.		

Table 188: Exploratory Data Analysis Engine: NFR-PORT-EDAE-001

ID	NFR-PORT-EDAE-001	Author	EXUS
Name	Dockerisation and integration with platform		
Description	The TANGO EDAE component should be dockerised.		
Priority	MEDIUM		
NFR Category	PORTABILITY		
Rationale	This will help to make the integration process easier.		

Table 189: Federated Learning Component: NFR-PER-FL-001

ID	NFR-PER-FL-001	Author	ATOS
Name	Federated Learning computing resources		
Description	The TANGO system must ensure that the training nodes in federated learning offer technical resources to perform the model training.		
Priority	HIGH		
NFR Category	PERFORMANCE		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	To make use of the Federated Learning component, several computing nodes are required in order to train the models based on the distributed data. The concrete technical requirements of these nodes might vary depending on the characteristics of the model and dataset to be trained.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	219 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 190: Federated Learning Component: NFR-PER-FL-002

ID	NFR-PER-FL-002	Author	ATOS
Name	Federated learning model performance definition		
Description	The TANGO system should provide the capability of defining thresholds for the model performance and training.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	Performance metrics could be provided in order to define the behaviour of the federated learning training process, including accuracy thresholds, model aggregation strategies, etc.		

Table 191: Federated Learning Component: NFR-SEC-FL-001

ID	NFR-SEC-FL-001	Author	ATOS
Name	Connectivity between training nodes and master node		
Description	The TANGO system must ensure connectivity between the training nodes and the master node in a federated learning scenario.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	In order to perform Federated Learning, the infrastructure where it is set up must guarantee the connectivity between the training nodes and the master node in order to allow the aggregation of the resultant training weights into the aggregated ML model.		

Table 192: MLOps Component: NFR-PORT-MLOPS-001

ID	NFR-PORT-MLOPS-001	Author	ATOS
Name	Virtualized deployment		
Description	The TANGO system must provide the MLOps component in a virtualized way.		
Priority	HIGH		
NFR Category	PORTABILITY		
Rationale	In order to ease the deployment of the component, it has to be designed and developed with its virtualization in mind. By doing it this way, the component can be easily deployed wherever, regardless the OS of the infrastructure of the server. This will be done using technologies such as Docker and/or Kubernetes which are widely used for this kind of purposes and have been proved to work properly in production environments.		

Table 193: MLOps Component: NFR-UI-MLOPS-001

ID	NFR-UI-MLOPS-001	Author	ATOS
----	------------------	--------	------

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	220 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Name	User Interface (UI)
Description	The TANGO system should provide an interactive UI where the MLOps component user can check the stat of all the ML experiments tracked as well as all the models or artifacts registered.
Priority	MEDIUM
NFR Category	USABILITY
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020, UR-DAR-M1-018
Rationale	In order to allow the end user a total control over what is being tracked with the MLOps component, a User Interface that shows all the experiments and artifacts registered should be provided with the component. Additionally, this UI might provide some functionalities to manage these experiments and artifacts, such as download or delete them.

Table 194: AutoML Component, NFR-PER-AUTOML-001

ID	NFR-PER-AUTOML-001	Author	ATOS
Name	AutoML computing resources		
Description	The TANGO system must ensure that the AutoML component is ran in devices that allow the execution of the training process of the models to be generated.		
Priority	HIGH		
NFR Category	PERFORMANCE		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	Deep Neural Network models require a high level of computing resources due to the number of operations required to train them. Although this computing resources consumption depends highly on the architecture of the model to be trained, the infrastructure where this component is to be deployed must provide the required resources in order to allow the training of several DNN models.		

Table 195: TornadoVM: NFR-COMP-TVM-001

ID	NFR-COMP-TVM-001	Author	UOM
Name	TornadoVM integration with a system software		
Description	The EDAE engine of the TANGO system could utilize the Java API of TornadoVM in Task 5.3.		
Priority	LOW		
NFR Category	COMPATIBILITY		
Rationale	The technology of TornadoVM implemented in Task 5.3 is based on Java. However, it can be operated along with code bases implemented in other programming languages, such as Python, via the interoperability with GraalVM and programming languages implemented on top of Truffle [53]. If the current NFR is not satisfied, the implementation of TornadoVM will stay in Java and it will be harnessed via a Python-binding implemented in GraalVM. Currently TornadoVM is compatible with GraalVM 22.3.2.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	221 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 196: Privacy Enhancing Component: NFR-COMP-PEC-001

ID	NFR-COMP-PEC-001	Author	UoG
Name	PEC-X_data_communication		
Description	The TANGO system must be able to initiate request/share data among other components to support data visualisation.		
Priority	HIGH		
NFR Category	COMPATIBILITY		
Rationale	PEC must be able to access/provide results to/from other TANGO components to provide a meaningful and holistic course of action (recommendation) to the users.		

Table 197 Privacy Enhancing Component: NFR-UI-PEC-001

ID	NFR-UI-PEC-001	Author	UoG
Name	PEC_graphical_user_interface		
Description	The TANGO system must be able to capture user input through usable interfaces to support the tasks of PEC.		
Priority	HIGH		
NFR Category	USABILITY		
Rationale	PEC must provide UIs to capture user input and display the outcomes and recommendations.		

Table 198: Privacy Assurance Tool: NFR-SEC-PAT-001

ID	NFR-SEC-PAT-001	Author	FUJ_LU
Name	Organization side interactions		
Description	The TANGO system must provide a UI to authenticate and request data in compliance to the GDPR.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-DMN-HT-003, UR-TDS-HT-011, UR-TUA-PA-002, UR-DUP-PA-006, UR-DMN-PA-009		
Rationale	Security: GDPR Compliance		

Table 199: Privacy Assurance Tool: NFR-UI-PAT-001

ID	NFR-UI-PAT-001	Author	FUJ_LU
Name	User side interactions		
Description	The TANGO system must provide an easy-to-use UI to authenticate and share data with preferences in a trustworthy and privacy aware way.		
Priority	HIGH		
NFR Category	USABILITY		
Related User Requirement	UR-DMN-HT-002, UR-TDS-HT-010, UR-TUA-PA-001, UR-DUP-PA-005, UR-DMN-PA-010, UR-DMN-PA-011		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	222 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Rationale	Usability: user-friendly interface and learnability.
-----------	--

Table 200: XAI Component: NFR-COMP-XAI-001

ID	NFR-COMP-XAI-001	Author	SQD
Name	Data communication		
Description	The TANGO XAI component must support communication over Hypertext Transfer Protocol (HTTP) with the rest of the system.		
Priority	HIGH		
NFR Category	COMPATIBILITY		
Related User Requirement	UR-DAR-M1-020, UR-DAR-BN-007		
Rationale	System will be exposed by an API.		

Table 201: XAI Component: NFR-PER-XAI-001

ID	NFR-PER-XAI-001	Author	SQD
Name	Optimal performance		
Description	The TANGO XAI component must have an optimal performance.		
Priority	HIGH		
NFR Category	PERFORMANCE		
Related User Requirement	UR-DAR-M1-020, UR-DAR-BN-007		
Rationale	Depending on the algorithm, the case and the volume of data, the component must have a proportionately optimal performance time and resource utilization.		

Table 202: XAI Component: NFR-EXIST-XAI-002

ID	NFR-PER-XAI-002	Author	SQD
Name	Hardware constraints		
Description	The TANGO XAI component must have the required resources available.		
Priority	HIGH		
NFR Category	HW & SW CONSTRAINTS		
Related User Requirement	UR-DAR-M1-020, UR-DAR-BN-007		
Rationale	Depending on the volume of the data and the complexity of the algorithms, there might be a need of a powerful Graphical Processing Unit (GPU).		

Table 203: XAI Component: NFR-PORT-XAI-001

ID	NFR-PORT-XAI-001	Author	SQD
Name	Dockerization and integration with platform		
Description	The TANGO XAI component should be dockerized.		
Priority	MEDIUM		
NFR Category	PORTABILITY		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	223 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Related User Requirement	UR-DAR-M1-020, UR-DAR-BN-007
Rationale	It helps make the integration process easier.

Table 204: XAI Component: NFR-REL-XAI-001

ID	NFR-REL-XAI-001	Author	SQD
Name	Reliable Explainability		
Description	The TANGO XAI component must produce reliable results.		
Priority	HIGH		
NFR Category	RELIABILITY		
Related User Requirement	UR-DAR-M1-020, UR-DAR-BN-007		
Rationale	The explainability report should be reliable so the model re-training is correctly facilitated.		

Table 205: RENOPS Component: NFR-MAIN-RENOPS-001

ID	NFR-MAIN-RENOPS-001	Author	XLAB
Name	Continuous integration and development		
Description	The TANGO RENOPS service should be developed in CI/CD pipeline.		
Priority	MEDIUM		
NFR Category	MAINTABILITY		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	The CI/CD pipeline should enable fast security patches and efficient integration of new features.		

Table 206: RENOPS Component: NFR-PORT-RENOPS-001

ID	NFR-PORT-RENOPS-001	Author	XLAB
Name	Containerized environment		
Description	The TANGO RENOPS service should be located inside a containerized environment.		
Priority	MEDIUM		
NFR Category	PORTABILITY		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	For service to be portable it should be located inside containerized environment. This will also enable the service to be more reliable and maintainable.		

Table 207: RENOPS Component: NFR-REL-RENOPS-001

ID	NFR-REL-RENOPS-001	Author	XLAB
Name	Orchestrator management		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	224 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Description	The TANGO RENOPS system could be managed by an orchestrator to scale the resources with demand.
Priority	LOW
NFR Category	RELIABILITY
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020
Rationale	To scale the resources according to demand, save resources when they are not needed and make operations more reliable, the service should be managed by an orchestrator.

Table 208: RENOPS Component: NFR-SEC-RENOPS-001

ID	NFR-SEC-RENOPS-001	Author	XLAB
Name	Secure communication		
Description	The TANGO RENOPS system must enable secure communication and maintain integrity.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-DAR-M1-012, UR-DAR-M1-020		
Rationale	For exchanged information, such as requested location and other user private data communication must be secure. Additionally, data must keep its integrity.		

D. System Requirements for Visualisation Techniques & User Interfaces

D.1 Functional Requirements

Table 209: User Interfaces: FR-S-UI-001

ID	FR-S-UI-001	Author	LST
Name	UI interactions		
Description	The TANGO UI must be able to receive messages from the system and notify the user accordingly, and vice versa.		
Priority	HIGH		
Technology Category	SUPPORT		
Rationale	The UI will be programmed to interact with the system through appropriate APIs, as well as with the user, through an interactive and responsive interface.		

Table 210: User Interfaces: FR-S-UI-002

ID	FR-S-UI-002	Author	LST
Name	Use of OPEN API specification		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	225 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Description	The TANGO technical offerings must follow the OPEN API specification for providing their data and parameters to the UI
Priority	HIGH
Technology Category	SUPPORT
Rationale	All the technical offerings need to be configured and monitored through the TANGO UI. A unified approach on having all the different technology offerings functionalities through the UI is needed, to achieve a homogeneous, extendible and user-friendly UI. Thus, Open API specification is required.

Table 211: User Interfaces, FR-S-UI-003

ID	FR-S-UI-003	Author	LST
Name	Input and output specification format		
Description	The TANGO technical offerings must use API REST back-end service that uses JSON documents for payloads. All the necessary input for the UI must use JSON documents for their payloads, in order for the UI to process them and visualize them under a common template. Specific requirements for these payloads will be drafted per technology offering, but all of them should provide an Authorization header for instance.		
Priority	HIGH		
Technology Category	SUPPORT		
Rationale	A unified way of passing all the parameters and data required in order to be presented through the UI is needed. The UI will handle the input from the various technology offerings in the same unified way, being agnostic of the internal processing of each offering. JSON format and the API REST specification is the most appropriate and well accepted one.		

Table 212: User Interfaces: FR-S-UI-004

ID	FR-S-UI-004	Author	LST
Name	Technology offerings specification and documentation generation.		
Description	The TANGO technology offerings must use the Swagger tool to define their interfaces, end points and the respective documentation. Each technology offering provider should document their interfaces and endpoints using Swagger. This will allow also easy testing and integration of their functionality.		
Priority	HIGH		
Technology Category	SUPPORT		
Rationale	The UI will be based on REST API calls that are handled through the endpoints of each technology offering. In order to have a well-defined, extendible and easy to test framework, the SWAGGER software will be used.		

Table 213: Visualisation and Reporting Techniques Component: FR-S-VIS-001

ID	FR-S-VIS-001	Author	SQD
----	--------------	--------	-----

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	226 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Name	Report visualizations from customer data
Description	The TANGO system must enable the hotel manager and the hotel group manager to visualize the reports produced by the analyzed data collected from customers.
Priority	HIGH
Technology Category	SUPPORT
Related User Requirement	UR-DAR-HT-009
Rationale	Hotel managers and hotel group managers should have a user-friendly visual representation of the reports.

Table 214: Visualisation and Reporting Techniques Component: FR-S-VIS-002

ID	FR-S-VIS-002	Author	SQD
Name	Visualized sustainability reports for clients		
Description	The TANGO system should be able to inform the clients about the hotel's sustainability data and environmental footprint in a user-friendly way.		
Priority	MEDIUM		
Technology Category	SUPPORT		
Related User Requirement	UR-DAR-HT-016		
Rationale	Clients should have a user-friendly way to view the sustainability reports.		

Table 215: Visualisation and Reporting Techniques Component: FR-S-VIS-003

ID	FR-S-VIS-003	Author	SQD
Name	Visualized sustainability reports for managers		
Description	The TANGO system should be able to inform the hotel management and the hotel association about their sustainability data and environmental footprint.		
Priority	MEDIUM		
Technology Category	SUPPORT		
Related User Requirement	UR-DAR-HT-017		
Rationale	Managers should have a user-friendly way to view the sustainability reports.		

Table 216: Visualisation and Reporting Techniques Component: FR-S-VIS-004

ID	FR-S-VIS-004	Author	SQD
Name	Visualized quality reports for customers		
Description	The TANGO system must be able to allow customers to visualize their own quality report only.		
Priority	HIGH		
Technology Category	SUPPORT		
Related User Requirement	UR-DAR-M1-009		
Rationale	Customers must have a user-friendly way to view their quality reports.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	227 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Table 217: Visualisation and Reporting Techniques Component: FR-S-VIS-005

ID	FR-S-VIS-005	Author	SQD
Name	Visualized quality reports for technical experts		
Description	The TANGO system must be able to allow the technical expert to visualize quality reports to help them improve the printer operation and the printing quality.		
Priority	HIGH		
Technology Category	SUPPORT		
Related User Requirement	UR-DAR-M1-010		
Rationale	Technical experts must have a user-friendly way to view quality reports.		

Table 218: Visualisation and Reporting Techniques Component: FR-S-VIS-006

ID	FR-S-VIS-006	Author	SQD
Name	Visualized reports for managers		
Description	The TANGO system must be able to produce visualizations for the reports by the analyzed costumer data.		
Priority	HIGH		
Technology Category	SUPPORT		
Related User Requirement	UR-DAR-M1-011, UR-DMN-RT-020		
Rationale	Managers must have a user-friendly way to view reports.		

Table 219: Visualisation and Reporting Techniques Component: FR-S-VIS-007

ID	FR-S-VIS-007	Author	SQD
Name	Alerts for unauthorized access.		
Description	The TANGO system must be able to show alerts in the device mapping dashboard when a potential unauthorized access to specific areas is detected.		
Priority	HIGH		
Technology Category	SUPPORT		
Related User Requirement	UR-PMT-M2-003		
Rationale	Alerts for unauthorized access must not only be produced but also be shown in a device monitoring dashboard.		

Table 220: Visualisation and Reporting Techniques Component: FR-S-VIS-008

ID	FR-S-VIS-008	Author	SQD
Name	Reports for unauthorized attempts.		
Description	The TANGO system should be able to visually show to supervisors reports of all unauthorized attempts in the platform system.		
Priority	MEDIUM		
Technology Category	SUPPORT		
Related User Requirement	UR-DAR-M2-009		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	228 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Rationale	The supervisor should have a user-friendly way to view reports of all unauthorised access attempts in the platform system.
-----------	--

Table 221: Visualisation and Reporting Techniques Component: FR-S-VIS-009

ID	FR-S-VIS-009	Author	SQD
Name	Data visualizations.		
Description	The TANGO system must allow specific users to visualize data.		
Priority	HIGH		
Technology Category	SUPPORT		
Related User Requirement	UR-DAR-RT-015, UR-DMN-RT-020		
Rationale	Users must be able to have a user-friendly way to visualize data.		

D.2 Non-Functional Requirements

Table 222: User Interfaces: NFR-PER-UI-001

ID	NFR-PER-UI-001	Author	LST
Name	Fast UI response		
Description	The TANGO UI must response fast to the user actions. Users should not wait long to see the options of the UI or to edit and save attributes. All their actions should be performed in a reasonable amount of time.		
Priority	HIGH		
NFR Category	PERFORMANCE		
Rationale	The response time of the UI is affecting the way that the users perceive the efficiency and functionality of the system. Slow UI reaction will make the system unusable and it will raise user's frustration, with the danger of abandoning its usage. Response time is a critical UI factor and it depends on the UI technologies used and the hosting platform specifications and resources.		

Table 223: User Interfaces: NFR-PORT-UI-001

ID	NFR-PORT-UI-001	Author	LST
Name	Responsive UI technology		
Description	The TANGO UI should be responsive in order to be able to efficiently work on the majority of the user devices.		
Priority	MEDIUM		
NFR Category	PORTABILITY		
Rationale	TANGO UI will follow a responsive design approach that will allow the easy navigation, information seeking, compatibility and usability of the TANGO offerings on various user devices without limiting functionality. This will affect also the performance, portability and maintainability of the UI.		

Table 224: User Interfaces: NFR-SEC-UI-001

ID	NFR-SEC-UI-001	Author	LST
Name	UI security		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	229 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Description	The TANGO UI must implement secure front-end technologies to avoid attacks. The technologies that will be used for implementing the UI must have inherent ways on protecting it from attacks. Furthermore, we should employ techniques that protect the UI from specific UI related attacks.
Priority	HIGH
NFR Category	SECURITY
Rationale	The UI is the point that the users interact with the system and allows its configuration and modification. Thus, it is one of the ways that attackers are trying to access the system and compromise it. For this reason, it is important that we will take measures against the related attacks.

Table 225: User Interfaces: NFR-UI-UI-001

ID	NFR-UI-UI-001	Author	LST
Name	The UI should be as simple and user-friendly as possible		
Description	The TANGO UI must provide an easily accessible, informative, user friendly and self-explanatory way for the user to access and configure the technical offerings and their attributes.		
Priority	HIGH		
NFR Category	USABILITY		
Rationale	The TANGO technical offerings are at some cases complex technologies that involve concepts that are not well known to the average end-user and they relate also to advanced security or data management terminology and concepts. The UI should allow the user to easily access, understand and use the technology offerings, being simple, self-explanatory and easy to use.		

Table 226: Visualisation and Reporting Techniques Component: NFR-COMP-VIS-001

ID	NFR-COMP-VIS-001	Author	SQD
Name	Data communication		
Description	The TANGO system must be able to initiate requests to the visualization component or send back responses to it with the appropriate data. Data sources must be accessible.		
Priority	HIGH		
NFR Category	COMPATIBILITY		
Rationale	The component will expose an API but also be programmed to initiate requests for data retrieval if necessary. The communication protocol will be HTTP and the data format will be in JSON.		

Table 227: Visualisation and Reporting Techniques Component, NFR-MAIN-VIS-001

ID	NFR-MAIN-VIS-001	Author	SQD
Name	Visualization component maintainability		
Description	The TANGO visualizations component must be maintainable.		
Priority	HIGH		
NFR Category	MAINTABILITY		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	230 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

Rationale	The maintainability of the visualizations component depends mainly on the quality of the source code. The data shapes or the data sources might change during the development.
-----------	--

Table 228: Visualisation and Reporting Techniques Component: NFR-PER-VIS-001

ID	NFR-PER-VIS-001	Author	SQD
Name	Visualizations performance		
Description	The TANGO visualizations component should produce the visuals in a timely manner.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		
Rationale	The data and the insights must be provided in a timely manner to ensure user experience.		

Table 229: Visualisation and Reporting Techniques Component, NFR-REL-VIS-001

ID	NFR-REL-VIS-001	Author	SQD
Name	Visualization's reliability		
Description	The TANGO visualizations component must present reliable results.		
Priority	HIGH		
NFR Category	RELIABILITY		
Rationale	The visualizations must correctly represent the analyzed report to the user.		

Table 230: Visualisation and Reporting Techniques Component, NFR-SEC-VIS-001

ID	NFR-SEC-VIS-001	Author	SQD
Name	Controlled access to visualizations		
Description	The TANGO visualizations component must be accessed only by certain users.		
Priority	HIGH		
NFR Category	SECURITY		
Related User Requirement	UR-DAR-HT-009, UR-DAR-HT-016, UR-DAR-HT-017, UR-DAR-M1-009, UR-DAR-M1-010, UR-DAR-M1-011, UR-PMT-M2-003, UR-DAR-M2-009, UR-DAR-RT-015		
Rationale	Not all users have access to all visualizations. This should depend on their perspective roles in the platform.		

Table 231: Visualisation and Reporting Techniques Component, NFR-UI-VIS-001

ID	NFR-UI-VIS-001	Author	SQD
Name	User friendly UI/UX		
Description	The TANGO system visualizations component should have user friendly UI/UX		
Priority	MEDIUM		
NFR Category	USABILITY		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	231 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final

Related User Requirement	UR-DAR-HT-009, UR-DAR-HT-016, UR-DAR-HT-017, UR-DAR-M1-009, UR-DAR-M1-010, UR-DAR-M1-011, UR-PMT-M2-003, UR-DAR-M2-009, UR-DAR-RT-015
Rationale	It should be easy for users to navigate and understand the visualized results.

E. System Requirements for the TANGO Integrated Platform

E.1 Non-Functional Requirements

Table 232: TANGO Integrated platform: NFR-COMP-TIP-001

ID	NFR-COMP-TIP-001	Author	INTRA
Name	TANGO components interoperability		
Description	The software components/services of the TANGO system must expose appropriate open and/or easily manageable and secure APIs (e.g., RESTful APIs over HTTPS) to ensure interoperability with other components/services		
Priority	HIGH		
NFR Category	COMPATIBILITY, SECURITY		
Rationale	In order to ensure the proper operation of the TANGO platform as an end-to-end system, supporting the use case scenarios of the different pilots, the interoperability among the different TANGO components/services must be implemented in an efficient and secure manner. To this end, it is envisioned that appropriate, widely accepted means of APIs shall be implemented.		

Table 233: TANGO Integrated platform: NFR-MAIN-TIP-001

ID	NFR-MAIN-TIP-001	Author	INTRA
Name	CI/CD best practices for software build, testing and deployment		
Description	The TANGO system must be connected with a continuous integration and continuous delivery framework which will ensure that all the components' developments can be smoothly integrated and deployed in an operational context.		
Priority	HIGH		
NFR Category	MAINTAINABILITY		
Rationale	Through the use of DevOps best practices accompanying the software development lifecycle, the integration of software updates in the TANGO system can be accelerated while ensuring a high degree of new bugs/issues prevention, through the execution of appropriate tests.		

Table 234: TANGO Integrated platform: NFR-MAIN-TIP-001

ID	NFR-MAIN-TIP-002	Author	INTRA
Name	Transparency of software updates to the operation of the TANGO platform		
Description	The updates on the components of the TANGO system should not cause any sort of interruption to the operation of the platform and remain transparent to the end user.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	232 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Priority	MEDIUM
NFR Category	MAINTAINABILITY
Rationale	This requirement aims to capture the possible need to run continuously and for large periods of time the different production environment clusters of services corresponding to the different pilots. In this context, it must be prevented to have any sort of service disruptions while making software updates.

Table 235: TANGO Integrated platform: NFR-PER-TIP-001

ID	NFR-PER-TIP-001	Author	INTRA
Name	Efficient software resources utilization		
Description	The software components of the TANGO system should be packaged and hosted in a way that will ensure efficient resource utilization.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		
Rationale	Considering the large amount of software components in TANGO, it will be required to use efficient means of software packaging to make sure that each component/application deployed at the TANGO integrated platform runs with only the minimum required dependencies from 3 rd party software/libraries and consequently hardware resources. This is also aligned with the overall requirement for energy efficiency.		

Table 236: TANGO Integrated platform: NFR-PER-TIP-002

ID	NFR-PER-TIP-002	Author	INTRA
Name	TANGO services scalability		
Description	The TANGO system should natively support the capability to easily scale the deployed applications according to the dynamic requirements.		
Priority	MEDIUM		
NFR Category	PERFORMANCE		
Rationale	This requirement accounts for any dynamically changing amount of requests towards specific applications/services of the TANGO system, instructing the corresponding application instances to scale up or down accordingly to accommodate the respective traffic requests. This is also aligned with energy efficiency in the sense of avoiding any underutilization of resources.		

Table 237: TANGO Integrated platform: NFR-PER-TIP-003

ID	NFR-PER-TIP-003	Author	INTRA
Name	Fast availability of software artifacts		
Description	The TANGO system could support the building of new artifacts (e.g., container images) within seconds.		
Priority	LOW		
NFR Category	PERFORMANCE		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	233 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Rationale	This is aligned with having short software development cycles, leading to quick identification of any issues (bugs) in the software and quick adoption of development updates.
-----------	--

Table 238: TANGO Integrated platform: NFR-PER-TIP-004

ID	NFR-PER-TIP-004	Author	INTRA
Name	Fast adoption of software updates		
Description	The TANGO system could support the components deployment from new artifacts within seconds.		
Priority	LOW		
NFR Category	PERFORMANCE		
Rationale	This is closely related to NFR 6 (Table 237), as another step of the software development cycle, following the software build procedures and leading to the quick adoption of development updates.		

Table 239: TANGO Integrated platform: NFR-PORT-TIP-001

ID	NFR-PORT-TIP-001	Author	INTRA
Name	Portability of software components and services across different hosting environments		
Description	The software components of the TANGO system must be packaged appropriately to be easily transferred and deployed to different hosting environments.		
Priority	HIGH		
NFR Category	PORTABILITY		
Rationale	This is important since it is expected that each component/service will need to be hosted in different environments (e.g., testing/staging vs production, pilot specific deployment environments vs core platform environment) during the project's duration. After the end of the project, it is important to have set the ground to easily exploit the software component developments across different hosting environments.		

Table 240: TANGO Integrated platform: NFR-PORT-TIP-002

ID	NFR-PORT-TIP-002	Author	INTRA, NOR
Name	Infrastructure vendor independence		
Description	The TANGO system could be able to run on any cloud or on-premise infrastructure.		
Priority	LOW		
NFR Category	PORTABILITY		
Rationale	Relying on a vendor-specific infrastructure feature would lead to vendor locking, therefore building TANGO system in a vendor-independent way would allow greater flexibility leading to better stakeholder/customer satisfaction.		

Table 241: TANGO Integrated platform: NFR-REL-TIP-001

ID	NFR-REL-TIP-001	Author	INTRA
----	-----------------	--------	-------

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	234 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Name	High availability and fault-tolerance of TANGO system
Description	The TANGO system must be highly available and fault-tolerant to avoid single points of failure
Priority	HIGH
NFR Category	RELIABILITY / MAINTAINABILITY
Rationale	Considering the production environment, it is important to be capable of quickly recovering from any software faults and to avoid any single points of failure, in order to minimize any service disruptions within the TANGO platform.

Table 242: TANGO Integrated platform: NFR-COMP-TIP-002

ID	NFR-COMP-TIP-002	Author	INTRA
Name	Deployment flexibility across different hosting environments of the TANGO system		
Description	The TANGO system should be able to support a level of automation in decision-making with respect to the deployment environment (cluster) where an application instance should be deployed, according to specific configuration requirements.		
Priority	MEDIUM		
NFR Category	COMPATIBILITY		
Rationale	This is a requirement of the TANGO architecture, regarding a central orchestration unit to decide about the deployment environment of specific applications/services, based on the configuration input.		

Table 243: TANGO Integrated platform: NFR-SEC-TIP-001

ID	NFR-SEC-TIP-001	Author	INTRA
Name	Secure, authenticated access to the TANGO integrated platform management interface		
Description	The management interface of the TANGO integrated platform must be accessible only for admin members and protected through access restrictions and authentication/authorization measures.		
Priority	HIGH		
NFR Category	SECURITY		
Rationale	This is a necessary requirement to protect the management of the TANGO deployment clusters from any unauthorized access.		

Table 244: TANGO Integrated platform: NFR-SUST-TIP-001

ID	NFR-SUST-TIP-001	Author	INTRA
Name	Energy efficient pilot executions		
Description	Each pilot use case must be using only the absolutely necessary components of the TANGO system to operate, eliminating transactions/processes that are not mandatory.		
Priority	HIGH		
NFR Category	SUSTAINABILITY		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment					Page:	235 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status:	Final

Rationale	This is a necessary requirement to make sure that no hardware resources/power consumption are wasted during the execution of the TANGO pilot use cases.
-----------	---

Table 245: TANGO Integrated platform: NFR-SUST-TIP-002

ID	NFR-SUST-TIP-002	Author	INTRA
Name	Energy efficient software deployments		
Description	The software component executables which are integrated in the TANGO platform must be lightweight to eliminate any redundant use of hardware resources and power consumption.		
Priority	HIGH		
NFR Category	SUSTAINABILITY		
Rationale	The software containerization approach can lead to energy-efficient software deployments, through minimizing the dependencies from 3 rd party software to the absolutely necessary ones and generating lightweight, standalone executables.		

F. Privacy, Ethics, Legal, and Societal Requirements

Table 246: Legal Requirement LEG-001

ID	LEG-001	Author	DBC
Name	Lawfulness of processing of the personal data		
Description	The TANGO system must ensure that the processing operations are taking place lawfully and according to the GDPR provisions.		
Priority	HIGH		
Category	PRIVACY / ETHICS / LEGAL COMPLIANCE		
Rationale	The processing of personal data should be based on a legal basis of processing according to article 6 GDPR. Because of the nature of the project and the proposed TANGO framework, the processing operations should be based on the legal basis of consent. On subsection 9.5.1 more details regarding the procedure of provision of consent from the data subjects are presented and those provisions must be followed to ensure GDPR compliance.		

Table 247: Legal Requirement LEG-002

ID	LEG-002	Author	DBC
Name	Exercise of the data protection related rights of the data subjects		
Description	The TANGO system must provide the users with the way that they could exercise their privacy related rights according to the GDPR.		
Priority	HIGH		
Category	PRIVACY / ETHICS / LEGAL COMPLIANCE		
Rationale	The TANGO platform shall provide the users – data subjects with a way to exercise their rights. This entails information in detail regarding the contact person		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	236 of 237
Reference:	D2.3	Dissemination:	PU	Version:	2.0
				Status:	Final

	with whom the data subject could communicate with and the way the request shall be applied.
--	---

Table 248: Legal Requirement LEG-003

ID	LEG-003	Author	DBC
Name	Privacy by design architecture		
Description	The TANGO system must be developed in compliance to the privacy by design related framework in terms of security and data privacy and integrity.		
Priority	HIGH		
Category	PRIVACY / ETHICS / LEGAL COMPLIANCE		
Rationale	Privacy and data security and integrity are considered to be a huge factor according to the GDPR. The TANGO platform should be developed in order for, by design, to adhere to those obligations.		

Table 249: Legal Requirement LEG-004

ID	LEG-004	Author	DBC
Name	Use of the AI technology		
Description	The TANGO system must implement the use of the AI with caution and in accordance with the related ethical provisions.		
Priority	HIGH		
Category	PRIVACY / ETHICS / SOCIAL / LEGAL COMPLIANCE		
Rationale	AI technology should be deemed lawful, ethical and robust.		

Table 250: Legal Requirement LEG-005

ID	LEG-005	Author	DBC
Name	Human participation		
Description	The TANGO system must follow the ethics recommendations regarding the participation of humans on the pilots of the project.		
Priority	HIGH		
Category	PRIVACY / ETHICS / SOCIAL / LEGAL COMPLIANCE		
Rationale	The participation of humans in the project must be voluntary and vulnerable people, such as patients or children, must be excluded.		

Document name:	D2.3 System Requirements and Specifications, Platform Architecture, and Privacy, Ethical, Social and Legal Impact Assessment			Page:	237 of 237	
Reference:	D2.3	Dissemination:	PU	Version:	2.0	Status: Final