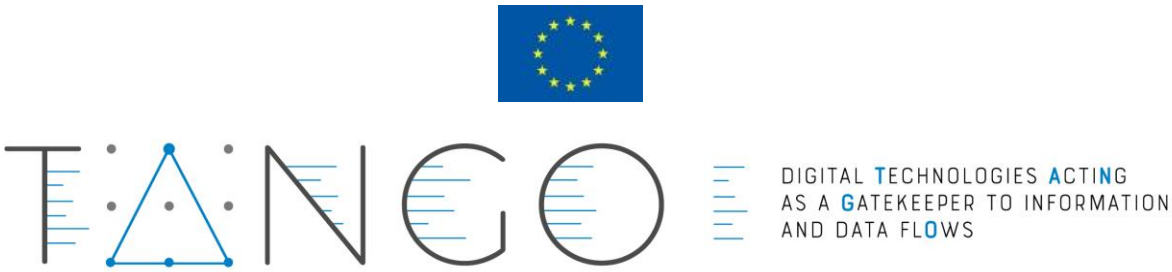

D1.3 – Data Management Plan and Research Ethics_final version

Document Identification			
Status	Final	**Due Date**	30/04/2025
Version	1.0	**Submission Date**	30/04/2025

Related WP	WP1	**Document Reference**	D1.3
Related Deliverable(s)	D1.2	**Dissemination Level**	PU
Lead Participants	KUL, DBC	**Lead Authors**	Alexandra Papageorgiou (KUL) Andrea Palumbo (KUL) Ioannis Drivas (DBC)
Contributors	ABI LAB, ACV, AHOP, ANYS, ATOS, CEA, CESGA, DBC, DUT, ECO. EGI, EXUS, FHG, FMAKE, FN, FSDE, FUJ_LU, IDIADA, INTRA, KUL, LC, MET, NORB, NTT, QBE, SQD, UMU, UoG, UoM, UPRC, UTH, VISAR, VTT, XLAB	**Reviewers**	Eilin Geraghty (ECO) Ville Ollikainen (VTT)

Keywords:
Data Management Plan, Research Ethics, Data Protection Policy, Responsible Research, and Innovation.

Disclaimer
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.

Document Information

List of Contributors	
Name	Partner
Sergio Salmerón	ATOS
Tomás Pariente, Ross Little, Iván Martínez	ATOS SP
Panos Matzakos, Eleni Veroni	INTRA
Pedro Pina, Pedro Cabral Santos, Marisa Brioso, Liliana Violante	SQUAD
Vitalii Demianets	NORBB
Ioannis Drivas	DBC
Anaxagoras Fotopoulos, Christina Nichiforov, Mohamed Khadmaoui	EXUS
Niklas Palaghias, George Sachpatzidis, Stefanos Panagiotou	QBE
John Zaras	SQD
Anja Božič, Nejc Bat	XLAB
Ioannis Koutsos, Chariton Palaiologk	FN
Dolores Ordóñez, Juan Ortells, Tayrne Butler	ANYS
Nicolas Belleville	CEA
Felix Hermsen	FHG
Ville Ollikainen	VTT
Kaitai Liang	DUT
María Hernández, Jesúss García, Antonio Skarmeta	UMU
Alexandra Papageorgiou, Andrea Palumbo	KU Leuven
Apostolos Apostolaras, Ilias Syrigos, Stergios Lantzos	UTH
Elena Avatangelou, Asterios Stroumpoulis	UPRC
Giulia Giussani	IDSA
Valeria Ardizzone	EGI
Lauresha Toska, Nelia Zinatullina, Peter Koller	ECO
Mpampis Chatzimallis, Francesco Mureddu	LIC
Raul Villalba	IDIADA
Anzellotti Emiliano, Marco Crabu, Piperno Piero	ABI
Dries Verhees	FMAKE
Miquel Martorell	CESGA
Juana María Serra	AHOP
Andreas Kopysov, Krystyna Ovchynnykova	VISA
Georgia Kyriakopoulou, Giorgios Tsiatiris	MET
Athanasios Stratikopoulos	UoM
Sakshyam Panda, Manos Panaousis	UoG
Elena Deimichei	NTT
Sven Rasmusen, Ibrahim Mecirdi, Moussa Ouedraogo	FUJ_LU
Jürgen Neises, Claudia Mertinger	FSDE
Evangelos Kotsifakos	ACV

Document History			
Version	Date	Change editors	Changes
0.1	18/04/2025	Alexandra Papageorgiou (KUL)	Document submitted for internal review
0.2	28/04/2025	Alexandra Papageorgiou (KUL)	Implementation of changes suggested by reviewers
1.0	29/04/2025	Alexandra Papageorgiou (KUL)	FINAL VERSION TO BE SUBMITTED

Quality Control		
Role	Name (Partner short name)	Approval Date
Deliverable leader	Alexandra Papageorgiou (KUL), Ioannis Drivas (KBC)	29/04/2025
Quality manager	Jürgen Neises (FSDE)	29/04/2025
Project Coordinator	Tomás Pariente Lobo (ATOS)	30/04/2025

Table of Contents

Document Information 2
Table of Contents 4
List of Acronyms 6
Executive Summary 7
1 Introduction 8
1.1 Project overview 8
1.2 Purpose of the document 8
1.3 Relation to other project work 8
1.4 Structure and methodology of the document 8
1.5 Disclaimers 9
2 TANGO Data Management Plan 10
2.1 Data summary 11
2.2 Data storage, quality, and security 13
2.3 Data availability and sharing between TANGO partners 14
2.4 Archiving, preservation, and deletion of data 15
2.5 Making data FAIR 15
2.5.1 Making data findable, including provisions for metadata 15
2.5.2 Making data accessible 16
2.5.3 Making data interoperable 17
2.5.4 Increasing data reuse 18
2.6 Other research outputs 19
2.7 Allocation of resources 20
2.8 Ethics 21
2.9 Other issues 22
3 TANGO Research Ethics and Compliance Protocol 23
3.1 The purpose of the GDPR and its core concepts 23
3.2 General principles of data protection and rights of the data subjects under the GDPR 23
3.3 Data protection policy 23
3.4 Data management and measures 24
3.5 Data protection impact assessment 24
3.6 Ethical issues and societal concerns in TANGO 24
3.6.1 General provisions 24
3.6.2 Research integrity 24
3.6.3 Ethics and data protection 24
3.7 Human Participation 25
3.8 Artificial Intelligence 26

4 Responsible Research and Innovation Practices 28
4.1 Open Science 28
4.2 Public Engagement 29
4.3 Ethics Integrity 29
4.4 Gender equality and inclusiveness 30
5 Conclusion 31
6 References 33
Annex I – General Data Information 34
Annex II – TANGO Data Summary 36
Annex III – TANGO Data Storage, Quality and Sharing between TANGO Partners 70
Annex IV – Data Availability and Sharing between TANGO Partners 78
Annex V – Archiving, Preservation and Deletion of TANGO Data 84
Annex VI – Making TANGO Data FAIR 87
Findable 87
Accessible 96
Data 96
Metadata 109
Repository 114
Interoperable 118
Reusable 124
Annex VII – Other Research Outputs 132
Annex VIII – Resource Allocation 136
Annex IX – Ethics 140
Annex X – Other Issues 143
Annex XI – EXUS dataset 145
Annex XII – NORB dataset 148
Annex XIII – Notice on Personal Data Processing and Providing of Consent Form 161
Annex XIV – TANGO Pilots Questionnaire 165

List of Acronyms

Abbreviation / acronym	Description
API	Application Programming Interface
AutoML	Automated Machine Learning
CP-ABE	Ciphertext-Policy Attribute-Based Encryption
CSV	Comma-Separated Values
DID	Decentralised Identifier
DMP	Data Management Plan
DPIA	Data Protection Impact Assessment
EDAE	Energy-efficient Data Analytics Environment
FL	Federated Learning
GDPR	General Data Protection Regulation
GUI	Graphical User Interface
HE	Horizon Europe
IPR	Intellectual Property Rights
ISM	Information Security Management
JSON	JavaScript Object Notation
JWT	JSON Web Token
ML	Machine Learning
MLOps	Machine Learning Operations
PAT	Privacy Assurance Tool
PEP	Policy Enforcement Point
Q&A	Questions and Answers
RRI	Responsible Research and Innovation
SMEs	Small and Medium-sized Enterprises
SP	Sticky Policies
TFEU	Treaty on the Functioning of the European Union
TSM	Trusted Service Manager
VC	Verifiable Credential
VP	Verifiable Presentation
WP	Work Package
ZKP	Zero-Knowledge Proof

Executive Summary

The present constitutes **deliverable D1.3** of the TANGO project, funded by the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101070052. It sets out the final version of the **Data Management Plan** and the **Research Ethics and Compliance Protocol** of the TANGO project and identifies relevant responsible research and innovation practices. It thus provides initial information about, among others, the data that is collected, generated and/or processed by the consortium within the remit of TANGO, and the manner in which such data and other research outputs are expected to be handled in terms of, *inter alia*, storage, security, and availability, while also laying out considerations related to making data and other research outputs findable, accessible, interoperable, and re-usable. Furthermore, the document touches upon ethical and legal requirements – including, but not limited to, those related to data protection – which were considered and followed by the consortium, along with the practices which guided research activities during the project.

This deliverable thus relates to all work packages of the TANGO project, and all the activities undertaken by the consortium during the project shall be performed by reference to this document.

Given that the information provided in deliverable D1.2 were preliminary and non-exhaustive and were set to be refined during the project lifecycle, this deliverable constitutes the updated version of the matters covered in the said deliverable.

1 Introduction

1.1 Project overview

The TANGO project aims at creating innovative solutions that will establish a stronger cross-sector data sharing in a citizen-centric, secure, trustworthy, and environmentally sustainable manner, promoting digitally enabled interactions across society, for people as well as for businesses. The overall TANGO environment will allow for user-friendly, secure, trustworthy, compliant, fair, transparent, accountable, and environmentally sustainable data management, having at its core technology components for distributed, privacy-preserving, and environmentally sustainable data collection, processing, analysis, sharing, and storage.

With partners comprising a balanced variety of renowned research institutes and universities, large, as well as small and medium-sized enterprises ('SMEs') that lead European research and development, and prestigious associations and think tanks, **TANGO will provide a high impact solution within the transport, e-commerce, finance, public administration, tourism, and industrial domains.** Through the provision of TANGO technologies, a trustworthy environment will be designed to act as a gatekeeper to information and data flows. Citizens and public/private organizations will be empowered to act and interact, providing data both online and offline.

More information about TANGO can be found on the dedicated website: www.tango-project.eu.

1.2 Purpose of the document

This document (deliverable D1.3) constitutes the final version of the Data Management Plan ('DMP') (led by KUL) and Research Ethics and Compliance Protocol ('Protocol') of the TANGO project (led by DBC). It also makes a final identification of relevant Responsible Research and Innovation ('RRI') practices (led by KUL). Overall, **the present document forms part of task T1.4 (Data Management, Research Ethics, and Legal Compliance)**, **falling under work package WP1 (Project and Technical Management).**

1.3 Relation to other project work

As previously discussed in D1.2, it should be reiterated that **the DMP and the Research Ethics and Compliance Protocol are related to all WPs of the TANGO project.** In other words, the activities in all WPs shall be performed in accordance with the considerations, principles, procedures, and guidelines laid out in the DMP and the Protocol. The same stands for the RRI practices identified, which are relevant for all research activities of the project.

1.4 Structure and methodology of the document

The structure of this document follows a logical, thematic progression designed to comprehensively address the TANGO project's obligations in terms of data management and research ethics. Following the introductory overview of the project's aims, accompanied by a detailed outline of the purpose, scope, and interrelation of the deliverable with other work packages (Section 1), **the core of the document is divided into two major components**: the **Data Management Plan ('DMP')** (Section 2) and the **Research Ethics and Compliance Protocol** (Section 3). Each component is developed using a combination of partner feedback, regulatory analysis, and best-practice guidelines provided by the European Commission ('EC') and international data stewardship frameworks. The methodology for compiling the DMP involved collecting structured responses from partners via questionnaires, complemented by a dedicated training session on FAIR data principles, and iterative consultations to ensure consistency and compliance. The ethical framework was informed by legal standards, notably the General Data Protection Regulation ('GDPR'), academic research ethics principles, and sector-specific insights drawn from TANGO's diverse pilot activities. The document concludes with an outline

of **Responsible Research and Innovation ('RRI')** practices (Section 4) and **annexes that provide further information on two sets of datasets used in TANGO** (Annex I – EXUS dataset as well as two templates: the **Notice on Personal Data Processing and Providing of Consent Form** for the cases where Pilot activities involve personal data processing (Annex III – Notice on Personal Data Processing and Providing of Consent Form), and the **TANGO Pilots Questionnaire** as circulated for the purposes of Work Package ('WP') 7 (Annex IV – TANGO Pilots Questionnaire). This approach ensures both operational clarity and strategic alignment with the broader goals of Horizon Europe. Finally, main take-aways are presented in the Conclusions (Section 5).

1.5 Disclaimers

The information provided in the DMP, the Protocol, the RRI practices part, and the Annexes does not constitute legal advice. Any user of this information uses it at their sole risk and liability.

2 TANGO Data Management Plan

Section 2 of D1.3 lays out the final version of the DMP of the TANGO project. The objective aimed at developing a DMP is to describe which research data are generated and used in the process of a research project, along with plans and strategies with respect to the organisation, documentation, storage, preservation, and sharing of that data, with respect to ethical and legal requirements applicable. Constructing a coherent DMP is crucial when it comes to facilitating data sharing, ensuring the sustainability and accessibility of data in the long-term and allowing data reuse for purposes of future research. Consequently, the TANGO DMP will set out how data collected, processed and/or generated during the project should be handled during its lifecycle and will refer to the procedures relevant to making data FAIR (*Findable, Accessible, Interoperable, Re-usable*). **As already indicated in the initial version of TANGO's DMP in D1.2, the DMP is a *living document***, iteratively adjusted, updated, and enriched as the project evolves, taking into account, among others, the generation or use of new data, changes to the original planning, changes in data/output access provisions or curation policies, or changes in consortium policies, practices, or composition.

The TANGO DMP is based on, and to a large extent reflects, the Horizon Europe ('HE') Data Management Plan Template, Version 1.1, dated 1 April 2022.[1] Other documents that were extensively studied and taken into consideration include guidance documents, information and/or good practices for the purposes of the preparation of the present DMP, including EC documents (such as the HE Program Guide) and third-party resources.[2] It goes without saying that fundamental legislative acts, both on an EU and national level were considered – including, but not limited to data protection and privacy legislation – shall guide data management during the TANGO project.

Information presented in this section derives from TANGO partners' individual responses to the 'TANGO Data Management Plan Questionnaire,' circulated on 10 January 2025, the answers to which are presented extensively in each of the respective sections below, along with further information provided by partners during the preparation of this deliverable.

Moreover, in order to ensure the completeness and robustness of the answers provided by partners to the said questionnaire, a 1-hour **DMP Training Session** was organised and took place online on **11 February 2025**. This session was co-organised with, and animated by, KU Leuven's Research Coordination Office, and consisted in a detailed presentation of Data Management in the context of Horizon Europe, along with tips and tricks for filing out the DMP questionnaire. It concluded with a Q&A part. More specifically, the presentation spanned across diverse topics, including: (i) the evolution of Open Science policies across Framework Programmes, covering projects within the period from 2008 to 2021; (ii) data management description; (iii) FAIR data principles and essentials; (iv) deposition of data in a trusted repository; (v) standardisation and documentation; (vi) metadata; and (vii) tools for management of active data.

This section is structured as follows: **Sub-section 2.1** provides the data summary – in other words, it describes the data handled by TANGO partners in carrying out their research activities under the project.

---

[1] European Commission, 'Horizon Europe Data Management Plan Template (Version 1.1)' <https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx> accessed 31 March 2025.

[2] Consortium of European Social Sciences Data Archives (CESSDA), 'Adapt Your Data Management Plan A List of Data Management Questions Based on the Expert Tour Guide on Data Management' (2019) <https://static-archive.cessda.eu/content/download/4302/48656/file/TTT_DO_DMPExpertGuide_v1.2.pdf,>; European Commission, 'Horizon Europe (HORIZON) HE Programme Guide (Version 4.1)' (2024) <https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf> accessed 31 March 2025; Science Europe, 'Practical Guide to the International Alignment of Research Data Management - Extended Edition' <https://zenodo.org/records/4915862> accessed 31 March 2025.

Sub-section 2.2 touches upon data access, storage, and security during the project. **Sub-section 2.3** discusses data availability and sharing between TANGO partners, while **Sub-section 2.4** pertains to the archiving, preservation, and deletion of data following the end of the project. **Sub-section 2.5** demonstrates how the FAIR principles are employed in the context of TANGO. **Sub-section 2.6** provides insights into other research outputs generated, and **Sub-section 2.7** does the same, but with respect to resource allocation. Finally, **Sub-section 2.8** discusses ethics issues, and **Sub-section 2.9** covers other questions of data management.

Each sub-section provides a summary of the input provided from partners, with the actual input provided by partners referenced to in separate Annexes (Annexes I-XII).

2.1 Data summary

This sub-section outlines the data/information handled by TANGO partners during the project, describing the data collected, generated, reused, and/or processed in research activities. **Table 1 in D1.2 included all the different categories of data which are handled in TANGO**,[3] along with different data types.

Bibliographic references that are purely employed to support an argument are not considered as research data for the purposes of Research Data Management ('RDM'). However, in the case where such sources constitute fundamental sources of information—and, as such, involve content analysis, the addition of annotation or coding, or extensive use of quotes—then bibliographic references are considered as research data. **This serves as a justification for the absence of datasets reported from partners whose contribution to TANGO was based on such references**. A general overview of the TANGO data is included in **Annex I – General Data Information.**

To the extent that the storage, sharing or other processing of personal data is involved, the fundamental data protection principles briefly outlined in the previous section should be respected. Further and final considerations with respect to the legal and ethical standards and principles TANGO partners shall comply with are provided in **Section 3** of this document.

The following first general conclusions can be drawn with respect to data handled across TANGO:

Across the TANGO project, **a wide variety of data is being collected, generated, or used by partners**. The types of data include numeric, textual, categorical, sensor data, images and video, electronic documents, metadata, and ML model outputs. These data are stored in formats such as CSV, JSON, XLS, HDF5, PDF, DOCX, JPEG, PNG, SVG, MP4, base64-encoded JWT, and NumPy arrays. The size of datasets varies significantly, ranging from just a few kilobytes for credentials or metadata, up to several terabytes for large machine data repositories like those used by FMAKE.

The time periods covered by these datasets generally align with the project's timeline, from 2023 to 2025. Some datasets are experimental or perpetual, such as AHOP's POI data, while others, such as identity verification or authentication data collected by QBE, are stored temporarily and deleted post-verification. Data like customer satisfaction surveys and transaction logs span months or years, depending on the use case.

Many datasets include personal data. These include banking data (ABI LAB & NTT), user identifiers (ACV), biometric or behavioural data (QBE), and contact information (FUJ_LU, LC). Others, such as EXUS (in its retail and manufacturing use cases), FMAKE, and several data science-focused tools (e.g., ATOS MLOps or AutoML) do not include personal data at all. Dummy data is also commonly used to simulate personal data, such as in VISAR's use case, ensuring compliance while enabling technical testing.

Regarding sensitive or special categories of data—such as racial or ethnic origin, political views, or union membership—these are not collected in TANGO. Portuguese law specifically forbids their

[3] See pages 13-15 of D1.2.

collection, and no data concerning children or vulnerable adults is processed in any of the documented use cases.

Pseudonymisation and anonymisation practices vary. Some partners pseudonymise sensitive attributes (e.g., ABI LAB & NTT, QBE, LC), while others anonymise video footage (e.g., IDIADA, FHG) or raw survey responses (e.g., CESGA). **A number of use cases operate on either synthetic data or encrypted/dummy data**, thus eliminating the need for traditional anonymisation or pseudonymisation. In cases where personal data must be identifiable—such as identity verification or issuing credentials—data is not anonymised due to the functional requirements.

Data is generally collected through direct interactions such as forms, sensors, online surveys, or project-specific apps. For example, ATOS's SSI agent relies on input from organisation or apps, CEA captures electromagnetic traces in a lab and MET aggregates POS data from its stores. The provenance of data is usually documented internally or through traceable APIs and software components.

Reuse of existing data is relatively rare but does occur. FMAKE reuses public 3D printing data from previous projects. UoM uses benchmark algorithms to test hardware acceleration. Most partners generate new data as needed for their specific tasks. In several cases, such as UMU or NORB, synthetic data or dummy data is created within the project's lifecycle to simulate real conditions without privacy risk.

Legal compliance and consent handling is a common theme. Most partners who deal with personal data have secured informed consent (e.g., ACV, FUJ_LU, QBE), with GDPR principles upheld. Tools like FUJ_LU's PAT platform offer detailed consent management capabilities. In other cases, like DUT or ATOS, the user retains full control over data upload, encryption, and deletion.

The data generated or collected across TANGO supports key project goals, including fraud detection, privacy risk scoring, identity verification, secure data exchange, machine learning training, and pilot evaluations. These activities map to a wide range of TANGO work packages, from WP2 on user requirements and WP3-5 on system design and AI to WP7 on pilot deployment and WP8 on dissemination.

Modifications to the datasets during the project include anonymisation, encryption, formatting changes, and the creation of model outputs. Some partners only handle raw or unmodified datasets (e.g., CESGA, ANYS), while others simulate transformations (e.g., NORB's layered encryption).

Dataset ownership and control are typically retained by the original collectors or data subjects. For example, financial data remains with the banks, hotel data with the hospitality organizations, and dissemination data with Lisbon Council. Some tools like the ABE toolset or the SSI Agent do not store data but instead process and return it in real time. Partners like ATOS and UMU provide components for secure processing but do not control the data itself.

Access to datasets is mostly restricted to specific TANGO partners involved in the relevant pilots or technical tasks. In most cases, access is limited to aggregated, anonymised, or encrypted formats, especially when personal data is involved. Public datasets like FMAKE's are exceptions, offering open access for broader use.

Processing activities include everything from identity verification, risk scoring, ML model training, encryption/decryption, and privacy enforcement. These are usually carried out by technical partners (e.g., ATOS, UMU, FHG, FSDE, IDIADA), while deployment and local data processing remain the responsibility of pilot organizations.

Data combination is not common within TANGO. Only a few use cases, like FUJ_LU's possible integration with UMU's PEP/PDP, explore the potential for merging different types of data (e.g., room sensors with guest preferences). Most partners treat their datasets as self-contained.

The utility of TANGO datasets outside the project is generally limited. While most datasets are tailored for project-specific goals, some exceptions exist. For example, FMAKE's 3D printing dataset

and CEA's side-channel data may benefit academic or industrial researchers. Lisbon Council's dissemination data could inform policymakers or institutions interested in secure data spaces.

Regarding international transfers, nearly all partners confirm that their datasets, especially those containing personal data, **will not be transferred outside the EU/EEA**. This ensures full alignment with GDPR and European data sovereignty principles.

Partners' individual input is provided under Annex II – TANGO Data Summary.

2.2 Data storage, quality, and security

During the TANGO project, **data storage practices varied significantly depending on the use case, the technical infrastructure of partners, and the nature of the data itself**. Many partners opted for **local storage**, particularly within their own facilities or pilot environments, such as ACV, AHOP, FUJ_LU, DUT, MET, UPRC, and FMAKE. Others, like ABI LAB & NTT, stored sensitive banking data within the secure infrastructure of the banks themselves. **A number of Pilots used cloud storage or hybrid solutions**. For example, UMU uploaded demonstrators to cloud data tools, LC used OwnCloud for backups, and QBE used Microsoft Azure in the Netherlands for T4.3. The Fides DRILL component, employed by NORB and VISAR, offered a distributed and legally compliant storage backend that integrates cloud technology. ATOS components such as the SSI Agent and ABE toolset only retained session data and not permanent user data, ensuring minimal storage risk. Some components, such as AutoML, MLOps, and FL, operated locally within partner environments, relying on their internal infrastructures to manage datasets and ML models. Certain entities like INTRA and CESGA reported no data storage responsibility, while others such as IDIADA and FHG coordinated shared local storage for the showcase phase across multiple institutions.

With respect to data backup strategies, practices varied widely as well. Some partners like CEA and MET implemented daily backups, while others such as LC scheduled weekly automated backups. UPRC maintained weekly backups locally, and VISAR performed automated backups before essential system updates or schema changes. ANYS ensured daily backups of persistence data through their NADIA platform, and QBE conducted continuous backups for seven days during certain processes. In contrast, several partners—particularly those who did not store data long-term, such as ATOS, DUT, IDIADA, FUJ_LU, and VTT (T4.1)—noted that backups were not applicable. Others, like ACV, indicated explicitly that they had no backup systems in place. For publicly available datasets, such as those hosted by FMAKE, local copies were maintained to safeguard accessibility.

To guarantee data integrity, quality, and confidentiality, partners implemented a variety of technical and procedural measures. ATOS ensured cryptographic signature proofs and attribute-based encryption (ABE) were in place for data used in identity verification and encryption/decryption processes. ANYS relied on an ETL process at the gateways to assure data quality. CEA secured data quality by optimally positioning electromagnetic probes and storing the output in controlled-access folders. DUT verified integrity via encryption and hash checking, while EXUS upheld strict data handling policies in line with its ISO:27001 certification, emphasizing limited access, MFA, and internal audits. UoG and FUJ_LU implemented robust validation checks and access controls, whereas UTH relied on a combination of input validation, user access authorization, and secure encryption to uphold quality and confidentiality. FHG, FSDE, and IDIADA deployed GDPR-compliant containers within secure environments, and VISAR's DRILL backend offered built-in integrity checks, validation processes, and restricted access for development teams. In general, where applicable, anonymization, encryption, or hashed identifiers were used to reinforce confidentiality.

Security policies and technologies adopted to protect data from unauthorised access or loss were equally comprehensive. Some partners, like EXUS and QBE, implemented full ISO:27001-aligned protocols, including firewalls, penetration testing, encryption, MFA, and VPN-based file transfers. ATOS tools operated over HTTPS with cryptographic validation, while NORB emphasized a multi-layered encryption architecture with randomized decryption mechanisms. UMU's SSI components ensured secure storage and controlled access, and DUT encrypted files using AES256 algorithms split

across chunks. FUJ_LU's Privacy Assurance Tool employed authentication protocols and encrypted storage, and LC maintained a combination of firewalls, antivirus software, and secure file transfers. In some cases—such as FMAKE, FSDE, and IDIADA—the data was considered non-sensitive, and as such, security was handled via internal IT infrastructure or by limiting persistence altogether. At UPRC, physical security measures were also reported, including on-site guards and disconnected hard drives for particularly sensitive data.

While some partners indicated no need or no application for data security or backup due to either the nature of the data or their specific role in the project (e.g., CESGA, INTRA, XLAB), **the majority took concrete steps to ensure that data under their control would be securely stored, regularly backed up, validated, and protected against loss or misuse.**

Partners' individual input is provided under Annex III – TANGO Data Storage, Quality and Sharing between TANGO Partners.

2.3 Data availability and sharing between TANGO partners

The same degree of variability amongst partners can be observed with regard to data access conditions, technical protocols, and mechanisms for ensuring availability, depending on the role of each partner, the nature of the datasets, and their responsibilities within the project.

In terms of access conditions, several partners provide either public datasets or share access under specific conditions. For example, CEA makes some side-channel datasets publicly available, while others are restricted to internal use. FMAKE and the University of Manchester (UoM) provide datasets that are openly accessible, with FMAKE datasets licensed under CC BY-NC-4.0. UMU, LC, and DUT share data internally in anonymized, aggregated, or pilot-specific formats, while EXUS, VISAR, NORB, and SQD limit access to curated or result-level datasets. Other partners, such as FUJ_LU, FHG, FSDE, and IDIADA, also restrict data sharing to output or showcase results only.

Conversely, a significant number of partners—such as ACV, CESGA, DBC, ECO, EGI, FN, INTRA, KUL, QBE, UoG, and VTT (T4.1)—**clearly state that no datasets are shared with other TANGO partners**. ABI LAB & NTT restrict access to participating banks and NTT Data; ANYS provides access via secure API tokens; and DUT limits access to users authenticated through the TANGO portal. In some cases, like NORB, access is limited to partners of specific pilots (banking, hospitality), while in others (e.g., ATOS), data availability is dependent on pilot-specific deployments.

From a technical perspective, most partners have implemented authentication and access control mechanisms. These include identity and access tools like Keycloak (FUJ_LU), Fiware Keyrock (ANYS), and MFA (QBE). EXUS uses its secure Nadia platform for hospitality data, while VISAR is developing its Fides backend with an API-driven query interface. DUT encrypts data using CP-ABE, with decryption keys protected and available only to authorized users. Standard formats (e.g., CSV, PDF, NumPy) and secure protocols (e.g., HTTPS, REST APIs) are common across multiple partners. In terms of record-keeping, FUJ_LU and UoG have detailed logging of access events, identity tracking, and data alteration history to ensure accountability and traceability.

To ensure data availability when needed by authorised users, several partners rely on internal cloud platforms, access-controlled repositories, or robust authorization mechanisms. CEA and UMU utilize OwnCloud for secure data sharing. LC distributes materials through secure internal channels. UTH ensures availability using a high-availability Kubernetes setup, while NORB uses its DnAMS tool for controlled access. ATOS, operating across multiple use cases (FL, MLOps, AutoML), makes data accessible at pilot sites upon deployment. In hospitality-specific applications, FUJ_LU provides data access through structured consent and access control workflows. For some partners, mechanisms are still under development or undefined, such as MET and VISAR, while others, like ACV, CESGA, DBC, ECO, and EGI, report no measures currently in place or that the topic is not applicable to their role in the project.

Partners' individual input is provided under Annex IV – Data Availability and Sharing between TANGO Partners.

2.4 Archiving, preservation, and deletion of data

Finally, at the conclusion of the TANGO project, partners have outlined diverse approaches to the management, retention, and deletion of data, based on the nature of the datasets, legal obligations, and institutional policies.

A large portion of partners—including ACV, FHG, FSDE, FUJ_LU, IDIADA, UoG, and NORB—**have confirmed that no data will be retained after the end of the project**, with most indicating that only dummy or test data was used throughout. Others such as ATOS (ID, SP, MLOps, AutoML, FL) will ensure that all configuration, session, and database data is deleted by the pilot administrators, using tools capable of overwriting virtual disk free space to prevent data remnants. In the case of EXUS, all project-related data will be securely deleted from cloud environments using ISO 27001 and SOC 2-compliant protocols, with audit trails and verification of the deletion process maintained.

Some partners will retain specific non-personal or public data after project completion. For instance, FMAKE expects data to remain available for up to three years, while UPRC will store data in its final project format for two years. LC commits to preserving public dissemination materials (e.g., newsletters, policy briefs) on the project website and in trusted repositories like Zenodo for at least five years, while ensuring that personal data is deleted in accordance with GDPR. Similarly, UMU and UoM will retain only examples or GitHub repositories, with synthetic data deleted within a short post-project timeframe (e.g., six months).

CEA will retain result graphs in CSV format while deleting unused datasets after it is determined they are no longer needed for analysis. MET plans to retain datasets in CSV format until the project's end, with no post-project retention indicated. UTH will delete pilot-related data upon completion of evaluations, and VISAR has used dummy data, so post-project handling is less critical. VTT's approach varies across tasks: T3.2 data, which contains no personal information, may be stored indefinitely, whereas T4.1 data will be completely deleted along with the associated pilot infrastructure.

In a few cases, such as SQD, the management and long-term storage of data falls under external control—in this case, RIAS—and *decisions will be updated later in the project*. QBE retains data in cloud storage only during the project (in JSON format), with deletion possible via API, DPO contact, or website request. Meanwhile, partners like DUT, CESGA, DBC, ECO, EGI, FN, INTRA, KUL, and XLAB have marked the issue as not applicable due to a lack of data handling responsibility.

In terms of backups, only a few partners explicitly mention retention periods. LC specifies that backups will be kept for five years and securely destroyed using GDPR-compliant methods, while most others either do not use backups or did not specify a backup strategy.

Overall, data retention strategies across the consortium are aligned with privacy, security, and legal compliance principles. Data will either be deleted securely at project end or retained in select cases where dissemination, public access, or further analysis is warranted and allowed.

Partners' individual input is provided under Annex V – Archiving, Preservation and Deletion of TANGO Data.

2.5 Making data FAIR

2.5.1 Making data findable, including provisions for metadata

Within the TANGO project, the use of persistent identifiers, metadata practices, and discoverability mechanisms show a mixed level of maturity and standardization across partners, reflecting the diversity of datasets, technological use cases, and partner roles.

Persistent Identifiers are used by a select group of partners. AHOP, ANYS, CEA, DUT, FMAKE, LC, and QBE confirmed their use of unique identifiers—ranging from transaction numbers to persistent object names, DOIs, or structured GitHub releases (UoM). DUT, for instance, names encrypted data using unique transaction numbers, while FMAKE embeds persistent labels within its HDF5 files. LC assigns DOIs to dissemination materials uploaded to repositories like Zenodo. In contrast, many partners (e.g., ABI LAB & NTT, FUJ_LU, IDIADA, FSDE, FHG) do not use persistent identifiers, often because the data originates from end-users and is not retained beyond project use. Several others (e.g., ATOS, CESGA, INTRA, KUL, XLAB) indicate this is not applicable due to the nature of their contributions.

Regarding the creation of rich metadata, a small group of partners has implemented structured approaches. CEA provides detailed technical metadata for side-channel measurements (e.g., sampling rates, implementation targets), while FMAKE includes labels in the dataset and documents its structure through a scientific publication. LC offers comprehensive metadata (titles, descriptions, authors, formats, versions) aligned with general standards. QBE and SQD also plan or provide metadata, with SQD noting the evolving nature of these definitions as work with RIAS progresses. However, many partners (e.g., ABI LAB & NTT, ACV, DUT, FUJ_LU, IDIADA, UMU, UoG) state that no metadata is created or needed due to limited data retention, while others (e.g., EXUS, NORB) defer the responsibility to data providers or indicate that FAIR compliance is not their decision.

When asked whether keywords are used for discovery a few partners, such as CEA, LC, and SQD, confirmed the inclusion of keywords to improve dataset discoverability. FMAKE notes that its dataset can be found via Google or through a linked publication, even though no keywords are embedded. Most others (including ABI LAB & NTT, DUT, FUJ_LU, and UoM) report that no such practices are in place or necessary.

As for naming conventions, approaches vary widely. LC employs clear, descriptive filenames that include the project acronym, WP number, content description, and versioning. CEA organizes data using a folder hierarchy that reflects implementation characteristics. FMAKE explains its naming convention publicly, based on printing layer indices. Some partners, such as ABI LAB & NTT and DUT, use system-generated feature identifiers or transaction numbers, while UMU follows naming conventions related to Verifiable Credentials and Decentralized Identifiers (DIDs). Most other partners (e.g., EXUS, ECO, VISAR) indicate this is not applicable or not defined.

The use of version control is confirmed by partners such as FMAKE, LC, SQD, UoM, and UPRC, who apply version numbers directly in file names or metadata. ATOS MLOps tracks versioning within MLFlow during machine learning experiments. Others, including DUT, QBE, ABI LAB & NTT, and FUJ_LU, state that versioning is not used, while many (e.g., IDIADA, CESGA, XLAB) deem it not applicable due to the absence of sharable datasets.

On the question of metadata harvesting and indexing (2.5.1.6), only CEA and LC explicitly provide metadata in formats that allow for discoverability—CEA does so via text files, while LC uses trusted repositories like Zenodo. FMAKE includes metadata within its dataset but notes it is not yet indexed. Others, including UMU, DUT, UoG, and ABI LAB & NTT, do not offer metadata in a harvestable form, and many more note that the issue is not applicable or has not been addressed.

In summary, while a number of TANGO partners have adopted robust metadata and data identification strategies—especially those working on dissemination, machine learning, or open science use cases—**many others do not apply persistent identifiers, metadata structuring, or version control due to the nature of their data, its limited use, or the lack of need for long-term accessibility.** As such, metadata maturity across the project is uneven, with best practices implemented selectively depending on use case relevance, openness of data, and long-term value.

2.5.2 Making data accessible

Only a limited number of partners—FMAKE, UoM, UPRC, and CEA (for one dataset)—**confirmed that at least some datasets will be openly available**. **Most other partners**, including ABI LAB & NTT, ACV, FHG, IDIADA, FUJ_LU, FSDE, NORB, QBE, and VISAR**, stated that data will not be**

shared, largely due to legal, privacy, or contractual limitations. Some, like LC, allow access to public materials while restricting personal or sensitive data due to GDPR compliance.

When it comes to restricted datasets, legal and contractual reasons dominate. Several partners—LC, QBE, NORB, FUJ_LU, FSDE, FHG, IDIADA, DUT—explicitly referred to data subject privacy, organizational ownership, or GDPR regulations as the primary barriers to sharing. A few datasets (e.g., ATOS SSI-related data and VTT T4.1) are restricted due to sensitive content like identity documents.

Data intended for public release—like those from FMAKE, CEA, and LC—**will be made available at the time of publication or shortly after internal validation, using platforms such as Zenodo or dedicated project repositories.** UoM provides access through GitHub and DockerHub, while UMU shares examples in technical documentation. Conversely, partners like EXUS, AHOP, and SQD clarify that they cannot decide on openness since data is externally owned.

Most accessible datasets will use standard protocols such as HTTPS or Zenodo APIs, while DUT relies on custom APIs and CP-ABE encryption. FMAKE supports HDF5 via open tools, and LC publishes in readable formats (e.g., PDF, MP4, CSV). UoG and others note access is tightly controlled and secured through authentication or institutional tools.

Where applicable, access control is enforced through internal systems or Pilot-specific mechanisms. CEA uses ciphered data transfers; LC enforces role-based permissions; NORB applies identity checks via eID or third-party providers; and UTH uses Influx DB authentication. However, most partners report that no access control or access committee is necessary because data is either not shared or inherently anonymous.

A small number of partners—FMAKE, LC, CEA, and UPRC—created openly accessible metadata. FMAKE provides metadata embedded in datasets and accompanied by example code on GitHub. LC ensures metadata is indexed and licensed through platforms like Zenodo. However, the majority—including ABI LAB & NTT, DUT, FUJ_LU, IDIADA, FHG, and others—do not create metadata or consider it unnecessary due to the restricted nature of their data.

Only a few partners provide metadata that directly enables data discovery or access. Where available, metadata may contain DOIs, format details, and licensing terms, though in many cases, either metadata is not generated or is tightly bound to internal-only data. LC guarantees metadata and dissemination material availability for at least five years post-project. FMAKE commits to long-term hosting via its institutional servers and public publications. CEA ensures metadata remains through associated publications even if datasets are removed. In contrast, most partners confirm that their data and metadata will only persist during the project lifecycle, with no guarantees for future access.

In terms of usability, only a few partners, such as FMAKE, UoM, LC, and CEA, **provide documentation or example code to facilitate data use**. UoM supports access through readme files and GitHub, while CEA uses common file types (e.g., .npy, .txt) and references standard libraries (e.g., NumPy). Most other partners do not require special tools due to either closed access or non-applicability.

Data is deposited by some partners into trusted repositories. LC, CEA, UPRC, FMAKE, and UMU have chosen repositories like Zenodo, GitHub, or institutional platforms. Zenodo provides DOIs for all uploads, ensuring persistent identification and resolution, as confirmed by LC and UPRC. For other partners, especially those handling private data (e.g., ABI LAB & NTT, NORB, QBE), no repository is used, and no identifiers are assigned.

2.5.3 Making data interoperable

In terms of data exchange and reusability, several partners have confirmed that their datasets are interoperable and adhere to widely accepted standards. For example, ANYS and CESGA use the FIWARE Context Broker, supporting NGSI data models to ensure interoperability. ATOS and UMU align with W3C standards such as Verifiable Credentials (VCs), Verifiable Presentations (VPs), and Decentralized Identifiers (DIDs), using JSON and JSON-LD formats to ensure compatibility across SSI platforms. FMAKE uses the HDF5 format, a common standard in data-intensive domains, while

FUJ_LU structures its data in standard JSON. LC ensures interoperability by using open formats like PDF, CSV, and MP4, while UoM, UPRC, and CEA confirm that their datasets are suitable for reuse and recombination.

Conversely, many partners either do not produce interoperable data or marked it as not applicable. For example, ABI LAB & NTT, DUT, NORB, UoG, and UTH state that their data is not interoperable due to encryption, proprietary formats, or use limitations. Several others—ACV, CESGA, IDIADA, KUL, EXUS, and others—did not provide input or deemed it not relevant due to the absence of primary data handling.

When it comes to data and metadata standards, those partners who work with interoperable datasets largely follow community or domain-endorsed practices. ATOS and UMU rely on the W3C Verifiable Credentials specification and JSON schemas, while FMAKE uses the HDF5 format. CEA supports its datasets with standardized vocabularies from the side-channel community and provides metadata along with publications. LC applies open, widely used formats and endorses interoperability best practices for reusability. SQD has a plan to support FAIR principles by preparing datasets in collaboration with RIAS, including assigning persistent identifiers and applying metadata standards.

Many partners did not adopt specific interoperability methodologies, with some noting that their data is either encrypted (e.g., DUT) **or not suitable for reuse** (e.g., NORB, FUJ_LU, UTH, UoG), and therefore, no interoperability methods are applied.

Regarding the use of project-specific ontologies or vocabularies, **most partners indicated that this scenario does not apply.** However, CEA commits to providing mappings to standard ontologies where necessary, and UMU notes it will publish mappings alongside relevant research. FMAKE documents its data structure embedded in the HDF5 format on its public website, while SQD plans to provide ontology mappings to commonly used standards based on user requirements. On the other hand, FUJ_LU, UoG, and UTH confirmed they do not use or intend to publish such mappings, typically due to the internal or closed nature of their data.

Lastly, with respect to whether datasets include qualified references to other data, the overwhelming majority of partners stated that they do not. Exceptions include LC, which references related TANGO datasets and previous research in its dissemination materials, and SQD, which aims to prepare curated datasets in cooperation with RIS for future reuse. Some partners, such as FHG, FSDE, and IDIADA, emphasized that although no external references are included in datasets, APIs and related documentation are available in the project's deliverables.

In summary, interoperability is actively pursued by a subset of TANGO partners who produce and manage data aligned with open formats and standards, especially in the context of federated learning, SSI, and smart systems. **However, for many others, the issue is either or not prioritised due to the nature of their involvement, the absence of raw data usage, or the use of encrypted, private, or non-reusable data.**

2.5.4 Increasing data reuse

Varying levels of maturity and commitment when it comes to supporting data reusability, documenting provenance, and implementing quality assurance are reported from TANGO partners. These factors are often shaped by the nature of the data, legal or institutional constraints, and the degree to which the data is meant for external use.

With respect to data documentation and validation, a limited group of partners—CEA, FMAKE, LC, UMU, UoG, UoM, and UTH—**have committed to providing "readme" files, guidelines, or technical documentation**. These documents describe data structures, methodologies, and use cases to support reuse and validation. FMAKE provides additional documentation via a scientific publication, and CEA includes instructions and makes results openly available through associated publications. UMU and ATOS also provide methodological details through TANGO deliverables and software

containers. Meanwhile, several partners, including ABI LAB & NTT, ACV, AHOP, and ANYS, do not provide such documentation, as data is not shared or not applicable in their context.

Only a few partners, notably FMAKE (CC BY-NC 4.0), UoM (Apache 2.0), LC, and UoG, **confirmed that their data or documentation is freely available under open licenses, facilitating broad reuse.** UMU plans to release examples and technical documentation under public deliverables or publications. Most other partners, including FUJ_LU, FHG, FSDE, IDIADA, QBE, and NORB, do not make their data publicly available due to privacy, legal, or organizational constraints.

Only a few partners foresee that their data will be reusable after the project. CEA and LC make data and dissemination materials accessible and usable by external stakeholders, including researchers and policymakers. UMU's technology modules (e.g., PEP, ID, SP) are intended as exemplars for future use. UoG and UoM confirm their contributions will remain available via GitHub repositories. The majority, however—including ABI LAB & NTT, FHG, FSDE, FUJ_LU, IDIADA, and others—state that their data will not be reused beyond the project or that reuse is not foreseen.

Provenance documentation is provided by a small group of partners. LC includes authorship, creation date, and versioning in its documentation, while CEA and FHG/FSDE/IDIADA provide such details through deliverables and project reporting. UMU commits to stating the data origin clearly. Most other partners either do not apply provenance standards or deem it unnecessary due to the limited or internal use of the data.

Only a few partners articulate structured quality assurance (QA) procedures. LC implements editorial reviews, version control, and internal approval workflows. CEA validates dataset integrity through hash comparisons between published and archived data. UTH outlines a plan involving statistical checks, legal compliance, and consistency checks. EXUS and FHG/IDIADA/FSDE refer to QA measures defined in TANGO deliverables. For many others, especially those not generating reusable data (e.g., ANYS, QBE, NORB), QA measures are not applicable or not defined.

In summary, while a subset of TANGO partners supports robust data reuse practices with open licensing, documentation, and metadata, the majority manage data under pilot-specific, restricted, or non-shareable conditions. Efforts to document provenance and ensure data quality are present in a few mature partners, but broader adoption would be required for full alignment with FAIR principles and long-term research impact.

Partners' individual input is provided under Annex VI – Making TANGO Data FAIR.

2.6 Other research outputs

In addition to data, a broad range of non-data research outputs, including software, workflows, models, protocols, publications, and communication materials, are being developed across the TANGO project. These outputs vary in scope and dissemination strategy, with some partners opting for open-access release under FAIR principles, while others retain certain outputs for internal use or as proprietary assets.

Many partners are generating digital research outputs, primarily in the form of software components, architectural frameworks, and machine learning models. For instance, ATOS is developing technical solutions like the SSI Agent, Federated Learning framework, and AutoML bundles. UMU and UTH are producing authentication software and privacy tools. DUT will release a standalone encryption-decryption plugin, while FHG, FSDE, and IDIADA highlight their contributions to the TSM module and its integration into the TANGO PEP. UoM is releasing a trained model through TornadoVM, and EXUS is developing the EDAE software, which will remain a proprietary part of their EXAI platform.

Several partners are also contributing architectures, models, and workflows, including SQD, which is active in defining TANGO's security models, threat modelling, and interface components. FN contributes a strategic sustainability roadmap, while LC focuses on communication assets—branding

kits, videos, policy briefs, and infographics—intended for public dissemination and stakeholder engagement. Others, like CEA, contribute software code, while partners like KUL and UPRC plan to produce scientific publications or white papers. VTT is preparing software (T4.1) and potentially academic publications (T3.2). A few partners, including FMAKE, note that physical outputs, such as printed objects, may be created as part of validation processes.

Several partners demonstrate a strong commitment to the FAIR principles in managing and disseminating their research outputs. For instance, LC ensures all digital media and policy outputs are made findable and accessible via platforms like Zenodo and the TANGO project website, using standard formats (e.g., PDF, PNG, MP4) and open licenses (e.g., Creative Commons). These outputs are also documented with metadata and explanatory materials to facilitate reuse and interoperability. UMU and ATOS share their outputs through open-access deliverables and code repositories (e.g., Git/GitHub), making architectural frameworks and technical documentation publicly accessible. SQD also emphasizes compliance with FAIR principles, with further specifications still under development. DUT will make its plugin freely available online, and UTH will upload its software to GitHub along with clear documentation. Some partners, such as CEA, manage and store their software internally using GitLab, while others—FSDE, FHG, IDIADA—highlight that their contributions are embedded in project deliverables or internal Docker containers. FMAKE complements its dataset with a scientific publication for validation and reuse, and FN intends to release strategic documentation in the project's shared space and open publication channels.

On the other hand, several partners, including EXUS, NORB, QBE, ANYS, AHOP, FUJ_LU, and others, do not plan to release software or digital assets, either because they are not generating them or due to their proprietary nature or internal purpose. EXUS, for instance, treats its EDAE tool as a trade secret and will not release it publicly.

In summary, TANGO partners are collectively producing a rich set of non-data research outputs. While some of these (especially software, models, and communication materials) are being released under open-access or FAIR-aligned practices, others remain proprietary, restricted to internal use, or tied to specific pilot implementations. The use of platforms like GitHub, Zenodo, and project-specific documentation ensures that key contributions from several partners will remain accessible and reusable beyond the life of the project.

Partners' individual input is provided under Annex VII – Other Research Outputs.

2.7 Allocation of resources

As far as the costs associated with making data and other research outputs FAIR are concerned, these are largely minimal or not applicable for the majority of partners. In many cases, partners are not generating data intended for public dissemination or reuse, or the outputs fall outside the FAIR scope due to legal, privacy, or operational reasons.

A small number of partners do identify specific FAIR-related costs. For instance, CEA reports these are embedded in its indirect project costs, while FMAKE incurs costs for hosting its public-facing website. LC outlines expenses related to repository fees, secure data storage, curation, and internal staff time, particularly from its communication team. FUJ_LU mentions that although its PAT system is inherently designed for secure data handling, updates would be needed to align with full FAIR compliance. NORB cites cloud infrastructure requirements for pilot data storage as a key cost component.

In terms of how these costs are managed, most partners who report FAIR-related costs plan to cover them through project budgets, internal funding, or work package allocations. For example, LC covers costs under its WP8 budget, and FMAKE, UMU, KUL, and UTH rely on project funding streams. CEA includes FAIR-related costs within its overarching indirect cost structure. FUJ_LU also highlights the potential for support through collaborations or external grants, while NORB relies on the pilot partner providing the required cloud services.

Regarding long-term preservation, a relatively small set of partners have strategies in place. CEA ensures preservation through data backup, publication in journals, and the use of open-access preprint archives, making results accessible beyond the project's lifecycle. LC applies a structured archival policy, storing public-facing materials in trusted repositories such as Zenodo and maintaining backups on internal systems for a minimum of five years post-project. FMAKE aims to retain data for several years after the project ends, while UMU and KUL will utilize public repositories to support availability. FSDE, IDIADA, and FHG defer to their institutional policies on data retention, which typically include secure short- to medium-term storage.

Other partners adopt more transient data lifecycles. FUJ_LU confirms that its data (via the PAT system) will be deleted at the end of the project, while UTH commits to retention for two months post-completion. NORB will also delete its data after pilot activities conclude. MET will store data on premise until project closure, with no further preservation defined.

Overall, while most partners report no significant FAIR-related costs, a select few are proactively managing the FAIRness and longevity of their outputs, particularly those contributing public datasets, software, or policy and communication deliverables. Where long-term access is intended, partners are making use of open-access platforms, internal infrastructure, and trusted repositories to ensure sustainability and compliance with FAIR principles.

Partners' individual input is provided under Annex VIII – Resource Allocation.

2.8 Ethics

The overwhelming majority of TANGO partners report that there are no foreseeable ethical or legal issues impacting data sharing. In contrast, **a few partners have identified specific ethical or legal frameworks guiding their work**. For example, ABI LAB & NTT indicate that banks independently manage legal and ethical obligations according to internal policies, with NTT Data not accessing any data directly. LC highlights that personal data collected (e.g., for events or newsletters) is subject to GDPR compliance, and that ethics considerations are documented in the Description of the Action (DoA). QBE acknowledges that the sharing of biometric data could raise legal concerns, while SQD explicitly references Portuguese legislation prohibiting the collection of sensitive personal data (e.g., political or religious affiliations), reinforcing strict GDPR compliance for all datasets curated with RIAS.

When it comes to informed consent for data sharing and long-term preservation, several partners confirm that informed consent protocols are in place. These include ACV, FHG, FSDE, FUJ_LU, IDIADA, LC, NORB, SQD, and UTH, who each ensure that participants are properly informed about how their data will be used, stored, and preserved. LC, for instance, includes clear terms regarding data storage and sharing in all consent forms. SQD follows GDPR-aligned procedures mandated under Portuguese law. UTH ensures user consent specifically for pilot evaluation monitoring. While UoG notes that informed consent is not currently applicable, it will be included if personalized data is collected.

On the other hand, partners such as ABI LAB & NTT, FMAKE, MET, and QBE confirm that personal data is either not shared or not included in their datasets, and hence consent for data reuse is not required. Others mark this question as not applicable due to the nature of their contributions or the absence of data collection.

In summary, while ethical and legal risks are minimal for most TANGO partners, those handling personal or sensitive data are taking the appropriate steps to ensure GDPR compliance, including the use of informed consent forms and alignment with national legal frameworks. In cases where data sharing is not foreseen or applicable, partners confirm that these obligations do not apply to their role in the project.

Partners' individual input is provided under Annex IX – Ethics.

2.9 Other issues

Across the TANGO project, the majority of partners report that they do not make use of national, funder, sectorial, or departmental data management procedures outside of those defined within the scope of the project itself. This applies to partners such as ABI LAB & NTT, ACV, AHOP, ANYS, ATOS, CEA, CESGA, FHG, FMAKE, FN, FSDE, FUJ_LU, IDIADA, INTRA, KUL, MET, NORB, QBE, UMU, UoG, UoM, UTH, VISAR, and XLAB. Some partners, including DBC, DUT, ECO, EGI, UPRC, and VTT (T3.2), consider the question not applicable based on their specific roles or non-involvement in data processing activities.

A small number of partners do adhere to external or institutional data governance frameworks. EXUS confirms compliance with the General Data Protection Regulation (GDPR) as well as adherence to the ISO 27001 information security standard, providing a structured approach to data protection and management within its organization. Similarly, LC follows a combination of internal data protection and IT security policies, Horizon Europe guidelines, and GDPR requirements to ensure consistent and responsible data management practices across its contributions to the project. SQD references national compliance measures under Portuguese data protection law, specifically noting adherence to the provisions outlined in Law No. 58/2019. This includes prohibitions on collecting sensitive personal data categories such as racial/ethnic origin, political or religious beliefs, and union memberships, which ensures the datasets curated with RIAS meet national and EU-level privacy standards.

In summary, while most TANGO partners manage their data within the project framework without invoking additional sectorial or national procedures, a few partners such as EXUS, LC, and SQD apply or plan to apply recognised standards and legal frameworks (e.g., GDPR, ISO 27001, national privacy legislation) to guide their data governance and security practices.

Partners' individual input is provided under Annex X – Other Issues.

3 TANGO Research Ethics and Compliance Protocol

Ethical compliance is considered a matter of outmost importance in the TANGO project in order to ensure that the developed framework is based on activities and procedures during which paramount ethical considerations were respected.

The primary focus of this section, in relation to compliance with ethical requirements, is the examination of personal data processing operations. For this reason, in the Deliverable D1.2, the main concepts, and the most important Articles of the GDPR have been thoroughly analysed, while the basic principles that rule the processing of personal data as well as the main rights that the GDPR provides to the data subjects have also been mentioned. Furthermore, the legal bases under which data is processed, the privacy policy which may be implemented by data controllers and the conditions for appointing a Data Protection Officer ('DPO') have already been presented. Special reference has been made to the manner in which the security of processing is ensured in the TANGO project, but also to the manner in which we deal with a personal data breach situation. Moreover, the definition of the DPIA has been analysed and the cases in which it should be performed in the TANGO project have been mentioned.

Additionally, **an ethical evaluation is being made regarding the human participation, especially for the third parties, in the project related activities**. This refers both to the procedures related to their actual participation in the project related activities and consequential aspects, such as the processing of their personal data.

Last but not least, **the development of and overall use of systems and models based on the artificial intelligence technology is envisaged in the TANGO project and is considered a factor that should be thoroughly assessed in terms of ethics.**

3.1 The purpose of the GDPR and its core concepts

To avoid repetition in this section, we refer to Deliverable D1.2.

3.2 General principles of data protection and rights of the data subjects under the GDPR

To avoid repetition in this section, we refer to Deliverable D1.2.

It is worth noticing that the process of onboarding users to the TANGO platform requires the usage of their passports in order to generate verifiable credentials to log into the system. Due to this, provisions have been made to respect GDPR and ensure the appropriate management of personal data:

- Integration phase: No personal data have been used. The components dealing with passport data used synthetic data instead. Therefore, GDPR related obligations did not apply in the integration environment.
- Pilots' environments: During the evaluation phase that will start in May 2025, the TANGO platform is deployed in each of the pilot's own infrastructure. Therefore, to ensure GDPR compliance, provisions have been made to sign a Data Processing Agreement (DPA) per pilot as well as to prepare and obtain signed Consent Forms from the participants prior to the evaluation. For each pilot, a specific partner has been named as Data Controller (usually the main pilot partner leader, or the associate technical partner), and partners whose components are involved in the data processing activities are named as Data Processors.

3.3 Data protection policy

To avoid repetition in this section, we refer to Deliverable D1.2.

3.4 Data management and measures

To avoid repetition in this section, we refer to Deliverable D1.2.

3.5 Data protection impact assessment

To avoid repetition in this section, we refer to Deliverable D1.2.

3.6 Ethical issues and societal concerns in TANGO

As already presented and analysed in detail in the D1.2, regarding all EU-funded activities, ethics plays a central role throughout the entire research process, from conception to completion. **Ensuring ethical compliance is not only a legal requirement but also a cornerstone of achieving true research excellence**. A thorough ethical assessment should begin at the proposal's conceptual stage, both to uphold the relevant legal standards and to improve the overall quality of the research. Conducting research ethically signifies applying fundamental ethical principles and relevant legislation across all research domains. **This evaluation is carried out through the Ethics Appraisal Procedure, a process designed to ensure that all ethical requirements are properly addressed and respected**.

3.6.1 General provisions

To avoid repetition in this section, we refer to Deliverable D1.2.

3.6.2 Research integrity

To avoid repetition in this section, we refer to Deliverable D1.2.

3.6.3 Ethics and data protection

This sub-section focuses on ethical issues related to data protection, a core aspect of research ethics and a fundamental right under EU law, as recognised by the Charter of Fundamental Rights and the Treaty on the Functioning of the European Union ('TFEU').

As already explained in D1.2, in research, data protection requires transparency with data subjects and responsible handling by organisations, including data minimisation, security, and timely deletion. Mishandling personal data can harm individuals and result in legal and reputational consequences for researchers. Notably, compliance with data protection laws does not automatically ensure ethical acceptability. Higher ethics risks arise when processing sensitive data, data from vulnerable groups, large-scale or invasive data operations, or when transferring data outside the EU. In such cases, a detailed ethics analysis of the methodology is essential, outlining all data processing activities, identifying associated risks, and explaining mitigation measures.

Furthermore, during data processing operations, obtaining explicit and informed consent from data subjects is essential. Project partners must clearly explain the purpose of research, participants' roles, and any potential risks before seeking consent. Consent must comply with the GDPR—freely given, specific, informed, and unambiguous—expressed through a clear affirmative act (written, electronic, or oral). Participants must be informed in plain language about their rights, the identity of the data controller, the purpose of processing, data retention, sharing practices, and how to withdraw consent.

When consent is the legal basis for processing, data must be used only for the specified purposes. Reusing data for new purposes without ensuring compatibility with the original intent may violate GDPR Article 5(1)(a), unless the new processing is demonstrably compatible with the initial one.

One of the key ethical challenges identified is the processing of personal data during the pilot phase of TANGO. This issue arises due to the involvement of third-party stakeholders who participate in the TANGO Pilots. It is our responsibility to ensure that all personal data is processed in full alignment with ethical principles and the requirements of the GDPR. To this end, dedicated consent forms have

been developed for each pilot. These forms are designed to provide participants with comprehensive information regarding the specific pilot activities, the types of personal data to be collected, the purpose of the data processing, the duration of data retention, and all other relevant details necessary for informed consent. In accordance with the GDPR provisions, this ensures that participants are fully aware of how their data will be used and are able to provide valid and freely given consent. A template of the consent form for the Tango pilots is included in Annex III, serving as a model for the consent forms used in all pilots.

Additionally, a targeted questionnaire has been developed for each pilot to gather detailed information and assess any potential risks or gaps related to data protection and ethical compliance. This process plays a critical role in safeguarding the privacy of individuals and ensuring that all project-related activities remain fully compliant with legal and ethical standards. **All partners involved in the pilots are required to complete the relevant sections of the questionnaire (see Annex IV) to support continuous monitoring and documentation of data processing activities.**

In the integration environment, we have decided to use synthetic data in order to ensure that only the personal data strictly necessary for the purposes of the project would be processed and no more than that. In the pilot Smart Hospitality, the partners have agreed to proceed by signing a Data Processing Agreement (DPA), as there will be multiple processors handling participants' personal data, alongside a controller who will maintain overall oversight and responsibility for the data processing activities. This is to ensure that all processing operations are carried out in accordance with the relevant legislative framework, particularly the GDPR.

To conclude, ensuring data security and integrity is ethically essential. Partners must implement appropriate technical and organisational measures, including anonymisation and pseudonymisation, to protect personal data. These measures should align with the GDPR's risk-based approach, adapting to the level of risk posed to data subjects' rights.

3.7 Human Participation

In this subsection, we outline the measures and procedures implemented to ensure full compliance with ethical standards concerning human participation—particularly third parties—in project-related activities. This includes both their direct participation in the pilots and the associated ethical considerations, such as the processing of their personal data.

As a starting point, **inclusion criteria have been established for participants in the TANGO pilots, where individuals will be involved in testing and evaluating the functionality of the TANGO framework**. These criteria are informed by the economic sectors identified as potential beneficiaries of the framework. Accordingly, individuals participating in each pilot are selected based on their relevance to those sectors. To ensure that the TANGO framework is assessed by domain-relevant stakeholders, participants will mainly be employees of the partner organisations responsible for the corresponding pilots.

The participation of individuals in all Project Tango pilots is strictly voluntary. To ensure that all participants are fully informed about the pilot activities and their involvement, a standardised template of an informed consent form has been developed (Annex 1). This document provides clear and detailed information regarding the purpose of the TANGO project and of the specific pilot, the nature and scope of participation and the participants' rights—including the right to withdraw at any time, without justification or penalty. Through this process, we can ensure that valid, explicit, and freely given consent is obtained from all participants, in line with ethical and legal requirements. Consent for participation will be obtained separately from consent for the processing of personal data, ensuring transparency and accountability.

In the case of the Smart Hospitality pilot, testing activities will take place at the premises of the pilot leader during the upcoming General Assembly meeting. Participants will be representatives from the project partner organisations attending the meeting. This approach ensures that the feedback and

evaluation of the framework will be provided by individuals with relevant experience, enhancing the quality and credibility of the pilot outcomes.

In the context of the Smart Manufacturing pilot of Project Tango, the participants involved in the testing activities will be employees of the participating organisation. This fact raises important ethical considerations, particularly due to the inherent power imbalance between employers and employees. It is therefore crucial to ensure that all ethical principles and requirements are fully respected and implemented.

First and foremost, it must be guaranteed that **employees' participation in the pilot activities is entirely voluntary and not perceived as an obligation or requirement linked to their professional duties**. To this end, specific consent forms have been developed (Annex III), clearly outlining the nature of the pilot activities and explicitly stating that participation is based solely on the individual's free choice, without any direct or indirect pressure from the employer.

Additionally, to further mitigate any potential ethical risks, participation has been limited to a small group of ten employees, all of whom hold managerial positions. This approach was adopted to ensure that participants are fully informed and aware of the pilot's objectives and processes, and to strengthen the assurance that their participation is truly voluntary.

Regarding the monitoring activities, which constitute part of the Smart Manufacturing pilot, it is important to highlight that **only corporate devices (work mobile phones) will be monitored within the context of the pilot**—under no circumstances will employees' personal devices be subject to monitoring. Moreover, the tracking of employees' location data will be performed in an approximate way, without using GPS technology, but rather through an alternative system that provides only general location information without precise accuracy.

Additionally, as part of our ethical oversight responsibilities**, we will conduct continuous monitoring of the pilot activities**. Specifically, anonymised questionnaires will be distributed to participants both at the beginning and at the end of the pilot. Through these questionnaires, participants will have the opportunity to express their views, experiences, and any potential concerns, ensuring that their participation remains voluntary and that no harm, stress, or discomfort is caused during the pilot. Participants are also explicitly informed that they may withdraw from the pilots at any time, freely and without any consequences.

In other words, a monitoring mechanism will be in place throughout the pilot, allowing us to intervene if any ethical concerns arise. Should any issues be identified that compromise the ethical standards of the project, particularly regarding the voluntary participation of the participants, we are committed to taking all necessary measures—including the suspension of pilot activities if needed—in order to ensure full compliance with ethical requirements.

Furthermore, **participant selection is strictly based on professional relevance, without any form of discrimination on the basis of gender, ethnicity, or background**. Vulnerable populations, such as minors or patients, are deliberately excluded from participation to avoid ethical and legal complications. As part of the ethical monitoring process, the anonymous questionnaires also include targeted questions related to inclusivity, fairness, and participant protection.

3.8 Artificial Intelligence

The TANGO platform is projected to have an overall positive social impact, having in mind mostly the privacy and environmental well-being preserving components, alongside the development of trustworthy AI.

The use of trustworthy AI, aligned with legal and regulatory frameworks, can have a positive social impact. Respect for human agency builds user confidence, ensuring secure and transparent interaction with AI systems. This enhances trust in digital services, as users understand how decisions are made and can request further information from developers. At the same time, trustworthy AI can reduce fears

related to job replacement, harm, or bias, increasing public acceptance and highlighting the supportive role of AI in daily life. Moreover, platforms like TANGO, by promoting environmental sustainability, can raise awareness, encourage responsible behaviour, and contribute to a better understanding of environmental issues.

The development and use of AI systems and models in the TANGO project is considered a key aspect that needs to be ethically assessed. The use of AI may present risks both during its development and implementation, particularly regarding compliance with legislation and human interaction. As part of this assessment, the project has considered key factors such as human agency, robustness, and safety, following the "Assessment List for Trustworthy AI (ALTAI)" provided by the European Commission.

Regarding human agency, the development of AI components is based on data directly provided by project partners. This ensures accuracy and relevance, improving the performance of the AI systems. In addition, users will be properly informed about their interaction with the AI system through visual indications, messages, and detailed documentation explaining the system's capabilities and decision-making process.

The risk of overreliance on AI has also been addressed. The systems developed will support users rather than make critical decisions independently. Transparency and explainability are enhanced through visualisations, confidence scores, and saliency maps, helping users understand the AI's operation and outputs.

In terms of robustness and security, **measures include continuous training with representative data, thorough testing scenarios covering edge cases, and data governance mechanisms to ensure data quality**. Monitoring tools will be used to detect risks and take corrective actions when necessary.

Privacy and environmental risks have also been assessed. No significant environmental impact has been identified, while all necessary procedures regarding data privacy are in place.

In conclusion, the TANGO project ensures ethical compliance in the development and use of its AI systems. Monitoring and evaluation will continue throughout the project to guarantee the development of trustworthy, transparent, and legally compliant AI solutions.

4 Responsible Research and Innovation Practices

As already outlined in Deliverable D1.2, the promotion of RRI is one of the main objectives of HE.[4] **This term broadly entails that the public (i.e., all relevant actors/stakeholders) is engaged in the research process to better align the goals and outcomes of research with societal needs and to address societal challenges**. RRI practices, reflected in the HE Regulation, the HE Program Guide, and the HE Strategic Plan (2021-2024),[5] relate to, for example, open science (including open access), public engagement, ethics and integrity, gender equality and inclusiveness, and societal impact.

In D1.2., the initial identification of RRI aspects to be considered by the TANGO consortium has been presented. An early overview of steps anticipated by TANGO partners has been provided, with further details to be included in the current deliverable, which constitutes an update of the DMP and research ethics.

TANGO aims to deliver societal impact by addressing data availability and sharing while respecting citizens' fundamental rights, particularly privacy and data protection. The project is focused on building trust, ensuring compliance, and promoting "green" data operations that benefit both organisations and citizens. TANGO can and is offering (in the ongoing pilots) privacy-preserving identity management solutions, enabling secure, user-centric data sharing.

The project follows a multidisciplinary approach, involving universities, research centres, think tanks, business associations, and diverse industries. These actors collaborate to ensure solutions meet societal values and needs. Additionally, TANGO will assess citizens' and organisations' acceptance of the solutions through six high impact use cases.

4.1 Open Science

For the open science practices that have been elaborated into the TANGO project, we also refer to section 4.1 of the Deliverable D1.2.

With respect to open science practices, regarding intellectual property rights (IPR) within the project, we have collected detailed input from all partners to ensure proper allocation of IPR. Key aspects identified included each partner's background contributions, their roles in project tasks, and the expected outcomes. **This process led to the creation of an IPR matrix in Excel format, structured in three parts.**

The first section of the matrix provided guidance on how to complete it. In the second section, partners listed the background knowledge or assets they contributed, specifying conditions of use and exploitation. The third section outlined expected exploitable results, identifying potential owners and contributors, links to project tasks, required protection (e.g., patents, licenses), exploitation potential, and whether each result was a key exploitable result.

The background items listed in the Consortium Agreement were included and reviewed by partners, who updated the list where needed. For exploitable results, responsible partners provided detailed information to support accurate IPR allocation.

[4] European Council, 'Council Decision (EU) 2021/764 of 10 May 2021 Establishing the Specific Programme Implementing Horizon Europe—the Framework Programme for Research and Innovation, and Repealing Decision 2013/743/EU' <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021D0764> accessed 18 April 2025.

[5] Directorate-General for Research and Innovation, European Commission, *Horizon Europe Strategic Plan 2021-2024* (Publications Office of the European Union 2021) <https://data.europa.eu/doi/10.2777/083753> accessed 18 April 2025.

After collecting all input, the consortium was able to identify and address any potential IPR-related conflicts. If conflicts arose, the designated partner took steps to resolve them by gathering more information, facilitating discussions between the involved parties and the coordinator, and seeking external advice if necessary.

4.2 Public Engagement

In line with the Horizon Europe objective of deepening the relationship between science and society through public engagement and inclusive research practices, **the TANGO consortium undertook a range of actions to foster public involvement and transparency.** Recognising the importance of this RRI principle, TANGO actively engaged with citizens and stakeholders throughout the project.[6]

TANGO's research activities and outcomes were communicated to the public via a dedicated website, social media, blog posts, newsletters, and scientific publications. The consortium also participated in and organised events to share project developments with the research community, industry stakeholders, and the wider public. These dissemination and communication efforts were structured according to the guidance of the initial Dissemination and Communication Plan (D8.1), developed under WP8/T8.1.

Throughout the project, TANGO partners continuously refined their public engagement strategy, ensuring that the needs and expectations of society were meaningfully addressed.

4.3 Ethics Integrity

The TANGO consortium recognises the critical role of ethics and integrity in research, ensuring that all activities comply with the highest ethical standards and legal requirements.[7] As already mentioned in D1.2, a dedicated ethics partner has been appointed, and an updated version of the research ethics and compliance protocol has been defined in section 3 above. Ethical considerations, including data protection and privacy, have been integrated into the project, with a comprehensive impact assessment addressing ethical, social, legal, and privacy issues as part of WP2/T2.6. Finally, certain ethical issues have been addressed under WP9, in which compliance with specific ethics requirements is ensured.

TANGO partners are committed to upholding the principles of research integrity as outlined in the European Code of Conduct.[8] This includes ensuring reliability, honesty, respect, and accountability throughout the research process. All researchers involved in the project are expected to follow established good research practices in line with the Code.

Finally, **with regard to human participation**, in the context of the Smart Manufacturing pilot of Project TANGO, as noted in section 3.2 of this deliverable**, the individuals involved in the pilot's testing activities will be employees of the participating organisation**. The employment relationship is widely acknowledged in legal theory and practice as one characterised by a structural power imbalance, whereby employers hold greater economic and organisational authority than employees. This asymmetry is evident in the employer's control over recruitment, remuneration, working conditions, and termination of employment. The inherent imbalance is also reflected in Article 28 of the Charter of

[6] European Council (n 5); Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013 2021 (OJ L 170, 12/5/2021) 1.

[7] European Commission (n 3).

[8] ALLEA - All European Academies, 'The European Code of Conduct for Research Integrity Revised Edition' (2017) <https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf> accessed 18 April 2025.

Fundamental Rights of the European Union, which guarantees workers the right to collective bargaining and action as a means to address and mitigate inequalities in labour relations.[9]

Having established the structural imbalance inherent in employment relationships, it is unsurprising that the participation of employees in the aforementioned pilot raises significant ethical considerations. It is therefore essential to ensure full compliance with all relevant ethical principles and requirements. To minimise potential ethical risks, participation has been limited to a small group of ten employees holding managerial positions, as previously stated.

Managing employees are those who perform management functions and often act on behalf of the employer within the organisation. They typically hold supervisory or representative roles, possess decision-making authority, and may influence or even determine company policy. Hence, the power imbalance between these executives and the employer is considerably less pronounced. Their consent to participate in the pilot can therefore be deemed genuinely voluntary, unlike that of other employees who might feel compelled to consent due to perceived obligations or fear of repercussions.

To further ensure voluntary and informed participation, anonymised questionnaires will be distributed to participants at both the beginning and the conclusion of the Pilot, as outlined in section 3.2. Through this approach, we aim to meet all necessary ethical safeguards and proceed with the pilot in a manner that respects participants' autonomy and rights.

4.4 Gender equality and inclusiveness

The TANGO consortium has effectively integrated gender equality principles into its work, aligning with the cross-cutting RRI practices in higher education and research.[10]

This includes ensuring a gender-balanced representation in co-design activities, such as workshops, and addressing gender considerations when relevant, such as in the training and evaluation of AI/ML algorithms. Additionally, to ensure the achievement of gender equality within the project, the research activities are being carried out by a gender-balanced team of researchers.

[9] Charter of Fundamental Rights of the European Union 2012 (C 326/392).

[10] European Council (n 5); European Commission (n 3); Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe—the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013.

5 Conclusion

The TANGO project, as detailed in Deliverable D1.3, has successfully established a robust and forward-thinking framework for data management, research ethics, and RRI practices. This document finalizes and refines the project's DMP and Research Ethics and Compliance Protocol, ensuring that all research activities within TANGO are aligned with European Union policies, notably the Horizon Europe framework and the General Data Protection Regulation (GDPR). Through systematic coordination among partners, detailed methodological planning, and regular training and consultation sessions, the project has effectively operationalized the principles of secure, ethical, and FAIR (Findable, Accessible, Interoperable, and Reusable) data management.

Throughout the project, data governance was treated with the utmost care, addressing a broad spectrum of data types ranging from personal information to technical metadata, sensor data, and synthetic datasets. The consortium applied high standards of data security, quality assurance, and access control, ensuring the integrity, confidentiality, and lawful sharing of data among partners. Where applicable, sensitive and personal data were handled in full compliance with GDPR principles, with dedicated mechanisms for informed consent, data minimization, pseudonymisation, and secure storage. The use of synthetic data in several pilot environments further exemplified the consortium's commitment to minimizing privacy risks while maintaining research efficacy.

In parallel, TANGO partners maintained a strong focus on research ethics. Procedures were established for the ethical recruitment and participation of individuals in pilot activities, ensuring voluntariness, transparency, and protection against any form of undue influence, particularly in workplace-based pilots. Dedicated consent templates, privacy notices, and ethical monitoring frameworks were created to uphold participant rights and project accountability. Furthermore, special attention was given to the responsible use of Artificial Intelligence (AI) within TANGO. Systems developed under the project were carefully evaluated against European Commission guidelines for trustworthy AI, emphasizing human oversight, transparency, robustness, and environmental sustainability.

The integration of Responsible Research and Innovation (RRI) practices into the project's core activities further enhanced the societal relevance and inclusiveness of TANGO's outputs. Open science initiatives, public engagement strategies, gender equality considerations, and ethical dissemination practices were incorporated wherever feasible, ensuring that the project's results can benefit a wide array of stakeholders beyond the consortium itself.

In terms of sustainability, the project took meaningful steps to ensure that key datasets, software tools, and communication outputs would remain accessible beyond the end of the project. Several partners committed to long-term archiving using trusted repositories such as Zenodo or institutional infrastructures, contributing to the broader scientific community and reinforcing the project's impact.

Most importantly, this deliverable is not a static artifact. It is designed to be updated in light of new findings, evolving stakeholder needs, and the dynamic realities of the project's technical implementation. **Feedback loops from pilot deployments, collaboration with external partners, and the integration of emerging standards will all contribute to the ongoing refinement of the DMP.**

As the TANGO project advances toward its objectives, this document will remain a central reference for guiding data-related decision-making and ensuring that innovation is aligned with ethical and societal values. It provides the foundation for a trustworthy, citizen-centric, and sustainable digital future, precisely the vision that TANGO seeks to realise.

Overall, D1.3 reflects the commitment of the TANGO consortium to uphold the highest standards of data stewardship, legal compliance, and research ethics. It ensures that the project's outcomes will not only fulfil contractual obligations but also contribute meaningfully to advancing trustworthy, secure,

citizen-centric data ecosystems in Europe. The project's legacy is thus built on a strong foundation of ethical integrity, technical excellence, and societal responsibility.

6 References

[1] ALLEA - All European Academies, 'The European Code of Conduct for Research Integrity Revised Edition' (2017) <https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf> accessed 18 April 2025

[2] Consortium of European Social Sciences Data Archives (CESSDA), 'Adapt Your Data Management Plan A List of Data Management Questions Based on the Expert Tour Guide on Data Management' (2019) <https://static-archive.cessda.eu/content/download/4302/48656/file/TTT_DO_DMPExpertGuide_v1.2.pdf,>

[3] Directorate-General for Research and Innovation, European Commission, *Horizon Europe Strategic Plan 2021-2024* (Publications Office of the European Union 2021) <https://data.europa.eu/doi/10.2777/083753> accessed 18 April 2025.

[4] European Commission, 'Horizon Europe Data Management Plan Template (Version 1.1)' <https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx> accessed 31 March 2025.

[5] ——, 'Horizon Europe (HORIZON) HE Programme Guide (Version 4.1)' (2024) <https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf> accessed 31 March 2025.

[6] European Council, 'Council Decision (EU) 2021/764 of 10 May 2021 Establishing the Specific Programme Implementing Horizon Europe – the Framework Programme for Research and Innovation, and Repealing Decision 2013/743/EU' <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021D0764> accessed 18 April 2025.

[7] Science Europe, 'Practical Guide to the International Alignment of Research Data Management - Extended Edition' <https://zenodo.org/records/4915862> accessed 31 March 2025.

[8] Charter of Fundamental Rights of the European Union 2012 (C 326/392).

[9] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) 2016 (OJ L 119/1, 4/5/16).

[10] Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013 2021 (OJ L 170, 12/5/2021) 1.

Annex I – General Data Information

Partner Name	Data details		Work Package and related Task
ABI LAB & NTT	Not applicable.		-
ACV	Dummy data.		WP7/T7.2
AHOP	Experiences POIS.		WP7/T7.2
ANYS	Room Condition.		WP7/T7.2
ATOS	ID	Identity data managed in project's IdM solution.	WP4, WP7
	SP	Sticky policies / CP-ABE.	WP4, WP7
	MLOps	MLOps support.	WP5
	AutoML	AutoML support.	WP5
	FL	Federated Learning support.	WP5
CEA	Combining Loop Shuffling and Code Polymorphism for Enhanced AES Side-Channel Security 10.5281/zenodo.10650736.		WP4
	Genelle Revisited 10.1109_LES.2024.3516853.		WP4
CESGA	Not applicable.		-
DBC	Not applicable.		-
DUT	Not applicable – individual processing of user data.		-
ECO	Not applicable.		-
EGI	Not applicable.		-
EXUS	Data from the Smart Hospitality, Smart Manufacturing, and Retailers' Pilots. *Please refer to Annex XI – EXUS dataset.*		WP5, WP6
FHG	Data from a track with the autonomous vehicle of IDIADA will be used on IDIADA site. Only scores calculated by IDIADA will be used by the Trustworthiness Scoring Module and the Privacy Risk scoring Module. Therefore, the modules itself do not use any collection of public or private data.		WP7
FMAKE	RAISE-LPBF-Laser data v1.1		WP7
FN	Not applicable.		-
FSDE	Data from a track with the autonomous vehicle of IDIADA will be used on IDIADA site. Only scores calculated by IDIADA will be used by the Trustworthiness Scoring Module.		WP7
FUJ_LU	Dummy data is collected from the pilot use case.		WP7/T7.2
IDIADA	Data from the vehicle will be used by FHG's Privacy Risk Scoring and FSDE's Trustworthiness Scoring Module.		WP7

INTRA	Not applicable.	
KUL	Not applicable.	-
LC	TANGO Dissemination, Communication, and Exploitation Dataset.	WP8
MET	Two datasets (Greece & Cyprus).	WP7/T7.7
NORB	Fides default corporate data dictionary.	WP7/T7.5
	Fides default individual data dictionary.	
	Please refer to Annex XII – NORB dataset	
QBE	Seamless user onboarding for users and devices: Quadible_internal_faces_01.	WP4/T4.2
	User continuous behavioural authentication: Quadible_internal_behavioural_01	WP4/T4.3
SQD	Assistance to the RIAS demonstrator with their Dataset Management, including but limited to preparing specific RIAS-TANGO Project datasets, in what regards EU Data legislation Compliance, meaning the separation of specific TANGO Datasets, the addition of persistent identifiers, the implementation of standard metadata access protocols, and, if necessary, the transformation of data for compliance with FAIR principles (present status to be updated further during the project).	WP6, WP7
UMU	**ID** Identity data managed in project's IdM solution.	WP4
	PEP Policy Enforcement Point.	WP4
	SP Sticky policies / CP-ABE.	WP4
UoG	Synthetic data is collected from the pilot use cases.	WP7/T7.2
UoM	TornadoVM: https://github.com/TANGO-EU-PROJECT/TornadoVM.	WP5
	docker- tornadovm: https://github.com/TANGO-EU-PROJECT/docker-tornadovm.	WP5
	DockerHub image: https://hub.docker.com/r/beehivelab/tornadovm-polyglot-graalpy-23.1.0-nvidia-opencl-container.	WP5
UPRC	Not applicable.	-
UTH	Not applicable.	-
VISAR	The dataset used in the pilot activities is based on personal data but is, in fact, dummy data. The dummy data is designed to resemble personal data.	-
VTT	Not applicable.	WP3/T3.2
	Passport data to be used for authentication purposes by SSI wallet.	WP4/T4.1
XLAB	Not applicable.	-

Annex II – TANGO Data Summary

1.1. Please describe the data you will collect/generate/use. In doing so, please refer to the **type** of data (e.g., numeric, textual, "electronic document", etc.), the **format** (the way in which the data is encoded for storage; you may also explain whether the format will change from the original to the processed/final data), the expected **size** (e.g. the storage space required and/or number of objects, files, etc.), etc.		
ABI LAB & NTT	Data used represent bank transactions provided by 2 collaborating institutes. The data are tabular data that contains numerical, categorical, text, and timestamp. The format will change since new features are created and all features are transformed and normalised before to be given as input to the model. We don't have details about the size of the data since we don't have direct access to it.	
ACV	Type of data: textual, numeric data. Format: Text Size: ≈ 1MB Data is encrypted in the database.	
AHOP	The data relating to Experiences and POIS is stored in xls format, with information relating to the location and segmentation of the various points of interest and experiences that can be offered to customers.	
ANYS	Nadia's datasets are stored in a MySQL database The collected data consists of sensor data, and all data is stored on AHOP premises. This information contains structures that store sensor information in string, numeric, bolean, etc. format. No personal data is stored or handled. Data can also be extracted via API.	
ATOS	**ID**	Data from the Onboarding issuer application (ePassport app) and Pilot organizations issuer application (Tango Dashboard) for issuing Verifiable Credentials ('VCs') is sent to a Self-Sovereign Identity ('SSI') Agent to issue the VCs to persons´ SSI Wallets, that are participating in the pilots. The format is specified by the W3C Verifiable Credentials Data Model. Each VC typically would be <<1MB (a few – tens of KB). For more detailed information, including non-normative examples of the VCs supported, please refer to section 3.1.1 of D4.2 Distributed Trust Management Framework.
	SP	Electronic files are sent to the ABE toolset component in a REST API call to be encrypted with their associated sticky policies, machine-readable format JSON. The size of electronic files used in the pilots is not known but expected to be MBs, while sticky policies would each be <<1MB (a few – tens of KB). The encrypted documents are returned to the pilot publisher organizations in response to the aforementioned request and not stored by the ABE toolset component. The reverse process occurs in the ABE toolset component to decrypt encrypted files by a pilot consumer organization providing the encrypted file and the decryption key. The ABE toolset component also creates the ABE public private keys for encrypting and decrypting the data upon request for the sticky policies. The ABE public private keys are stored in the ABE toolset component and will be a few KBs. Public keys are made publicly available for encrypting the data. For more detailed information, please refer to section 3.3.3 of D3.2 Distributed Privacy-preserving Data Management and Storage.

1.1. Please describe the data you will collect/generate/use. In doing so, please refer to the **type** of data (e.g., numeric, textual, "electronic document", etc.), the **format** (the way in which the data is encoded for storage; you may also explain whether the format will change from the original to the processed/final data), the expected **size** (e.g. the storage space required and/or number of objects, files, etc.), etc.		
	MLOps	MLOps support bundles together with several open-source software tools (MLFlow, MinIO, Jupyter Lab). MLFlow stores metadata about artifacts in MinIO (e.g., datasets, ML models versions, and metrics of experiments training and executing those models). This is done in the format required by MLFlow. The components are meant to be deployed at the customer side; therefore all the data is the responsibility of the pilots where the tool is deployed.
	Auto ML	AutoML aims at facilitating the initial steps of training ML Models. The component is based on the MLJar library and complies with its requirements for data. In particular, the component expects a dataset in .csv format to initiate the training. The component is meant to be deployed at the customer side (e.g., a data scientist laptop), therefore all the data management is the responsibility of the pilots where the tool is deployed.
	FL	FL helps training a ML model using data locally in different disperse nodes rather than a single repository or data lake. The client nodes of the FL network are deployed locally with access to the data on that node, while a "central" server node aggregates the weights of the locally trained models. Therefore, the data is accessed only at the customer side, while the central node is only aware of the locally trained models and the final aggregated one. Data can have several formats (e.g., in the case of the FMAKE pilot a HDF5 format was available at the local nodes).
CEA	For both datasets, data are electromagnetic traces, stored as NumPy files. First dataset is 3.1GB large. Second dataset is 11.9GB large.	
CESGA	Data related to customer satisfaction is provided in xls format; these files contain data related to customer opinions and ratings with data type, string, and numeric.	
DBC	Not applicable.	
DUT	Data to use: the design file of FMAKE smart manufacturing use case. Data to generate: the encrypted data of design file. Type: electronic document, format is user-customized, which cannot be opened with text reader. The format will be changed after encryption (to unreadable data) and recovered after decryption.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	**Smart hospitality**	Numeric, sensor name. Size: 230 MB Format JSON/CSV
	Smart manufacturing	Numeric, image. Size: 60 GB used, in total 784 GB. Format: hdf5

1.1. Please describe the data you will collect/generate/use. In doing so, please refer to the **type** of data (e.g., numeric, textual, "electronic document", etc.), the **format** (the way in which the data is encoded for storage; you may also explain whether the format will change from the original to the processed/final data), the expected **size** (e.g. the storage space required and/or number of objects, files, etc.), etc.		
	Retail	String, numeric. Size: 21 MB Format: CSV
FHG	A test dataset of a 30min ride has been generated by IDIADA and shared with Fujitsu and FHG. The test dataset is used to generate trustworthiness scores by FHG's Privacy Risk Scoring and Fujitsu's TSM. For a configurable window of time (up to 1000 entries), scores on attributes will be submitted to the TSM by IDIADA in a JSON format. The scores are floating point numbers in the range of 0.0 to 1.0.	
FMAKE	The data has been collected during a previous project. Type = machine data Format = HDF5 Expected size = 1 TB	
FN	Not applicable.	
FSDE	A test dataset of a 30min ride has been generated by IDIADA and shared with Fujitsu FSDE and FHG. The test dataset is used to generate trustworthiness scores by FHG's Privacy Risk Scoring and Fujitsu's TSM. For a configurable window of time (up to 1000 entries), scores on attributes will be submitted to the TSM by IDIADA in a JSON format. The scores are floating point numbers in the range of 0.0 to 1.0.	
FUJ_LU	Type of data: textual, numeric data. Format: JSON Size: ≈ 1MB Data is encrypted in the database.	
IDIADA	The test dataset is used to generate trustworthiness scores by FHG's Privacy Risk Scoring and Fujitsu's TSM. For a configurable window of time (up to 1000 entries), scores on attributes will be submitted to the TSM by IDIADA in a JSON format. The scores are floating point numbers in the range of 0.0 to 1.0.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	LC collects and generates textual data, such as press releases, newsletters, policy briefs, social media posts, branding guidelines, and reports. It also generates visual content including logos, infographics, brochures, videos, and animations. The formats include PDF, DOCX, HTML, PNG, JPEG, SVG, MP4, and CSV. Metadata related to dissemination activities, such as event dates, communication channels, and engagement metrics, are also collected. The dataset size is estimated between 5GB and 10GB over the course of the project.	
MET	Type of data: All types (numeric, alphanumeric, Boolean, datetime, etc.) Format: CSV Size: Unknown	
NORB	Numeric, text, boolean, single_select, double_select. *Please refer to Annex II for more detailed information on NORB datasets.*	

1.1. Please describe the data you will collect/generate/use. In doing so, please refer to the **type** of data (e.g., numeric, textual, "electronic document", etc.), the **format** (the way in which the data is encoded for storage; you may also explain whether the format will change from the original to the processed/final data), the expected **size** (e.g. the storage space required and/or number of objects, files, etc.), etc.		
QBE	**T4.2**	Seamless user onboarding for users and devices. Data from identity document of the user, including a facial image of the user. The data is temporary stored only during the verification process; once the verification is completed, the data is deleted. The output is the verified information of the identity document which is sent to the SSI Issuer (T4.1).
	T4.3	User continuous behavioural authentication. Data from accelerometer, gyroscope, magnetic field sensors, touch events, typing events, face images, location, Bluetooth, and WiFi traces. The aforementioned data (except the face images which are not stored) are stored in JSON format. The final data is the output of the authentication mechanism, indicating if the legitimate user is accessing the target app or an imposter.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS still under definition (present status to be updated further during the project).	
UMU	**ID**	Mainly electronic documents that certify and contain the data (e.g. Verifiable Credentials (VCs) and Verifiable Presentations (VPs), usually in JSON-LD format (e.g., following W3C's Verifiable Credential format). In most cases they will not need to be stored after demonstrator scenario execution ends, except for examples/demonstrations. In those cases, each would be <<1MB (a few – tens of KB).
	PEP	Electronic documents representing the policies, machine-readable format JSON. Documents would be <<1MB (a few – tens of KB) each and would not need to be stored except for a few example documents for reference (plus documentation on them, e.g. through deliverables).
	SP	Same as above.
UoG	Type of data: textual, numeric data. Format: JSON Size: ≈ 1MB Access to data is done only through authorization (role-based access).	
UoM	The technology developed by UoM, TornadoVM, interfaces with no-proprietary source code that is open (permissive licenses). In our case, we do not collect any data from use cases. In the contrast, we are implementing hardware accelerated algorithms that will be open-source available in GitHub (TANGO fork of TornadoVM repository). Then, the TornadoVM technology will generate at runtime the accelerated artifacts (OpenCL kernels – type: source code), which will be indirectly used by the EDAE (T5.1). The storage size of the generated files is in the range of 1-2 kB. UoM technology will also utilize the open source tornadovm benchmarks (available in the TornadoVM GitHub repositories) in order to collect energy metrics and train a ML model. The energy metrics are stored in text files and the size of them is around 2-3 KBs per hardware device (e.g., GPU, CPU). This workflow is containerized and the source code of the Dockerfile is available in GitHub (TANGO fork of docker-tornadovm repository). The docker container is	

1.1. Please describe the data you will collect/generate/use. In doing so, please refer to the **type** of data (e.g., numeric, textual, "electronic document", etc.), the **format** (the way in which the data is encoded for storage; you may also explain whether the format will change from the original to the processed/final data), the expected **size** (e.g. the storage space required and/or number of objects, files, etc.), etc.		
	available in Docker-Hub, and the size of the storage space of the image is 5.8-6 GB.	
UPRC	During the first months of the project, the data collected by UPRC were based on questionnaires that representatives from the pilot users answered during the user requirements collection process of WP2. The data described the existing processes of the pilot users and other information related to their use cases. Data is in electronic format. They are mostly free text descriptions. Moreover, during the last months of the project, UPRC is going to collect feedback from the pilots in order to evaluate the components and the TANGO platform. UPRC will not collect any personal or other sensitive data.	
UTH	The data will refer to a mobile device operational parameter such as performance metrics (e.g., power consumption, throughput, latency, access to specific networks, nearby base stations, or access point lists). Data is comprised of log files, which for a single hour corresponds to 150 Kbytes.	
VISAR	The data used in the pilot activities is dummy data. This dummy data is used to simulate the types of personal information that might be collected in a real case scenario.	
VTT	**T3.2**	Not applicable.
	T4.1	Electronic data describing passports of users, stored as base64-encoded JWT verifiable credentials. At most 100 documents, each at most 1KB. Used for providing secure identification services. The data contents are not analysed for research uses.
XLAB	Not applicable.	

1.2. Please indicate the time period covered by the dataset.	
ABI LAB & NTT	The data extraction period for fraudulent transactions has been set from March 2023 to December 2023, covering a total of **9 months**.
ACV	The UI component stores the data for the time period that the Pilot will run.
AHOP	POI data are perpetual. Those related to Experiences correspond to those published in 2025.
ANYS	NADIA historical data are discontinuous from July 2024, being every 5 minutes from 10 March 2025 onwards.
ATOS	**All components** cover data throughout the project.
CEA	October 2023 to February 2024 for the first dataset, and July 2024 for the second dataset.
CESGA	Customer satisfaction data covers January 2021 through December 2024. The complete dataset is expected to be available at least until March 2025.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.

1.2. Please indicate the time period covered by the dataset.		
EGI	Not applicable.	
EXUS	**Smart hospitality**	July and October 2024.
	Smart manufacturing	Not applicable.
	Retail	1 March 2023 to 1 September 2023.
FHG	The test dataset includes a recording of 5 minutes CAN data, about 320,000 lines, and an anonymized 30min video.	
FMAKE	Data was generated a few years ago.	
FN	Not applicable.	
FSDE	The test dataset includes a recording of 5 minutes CAN data, about 320,000 lines (FSDE), and an anonymized 30min video (FHG).	
FUJ_LU	The Privacy Assurance Tool (PAT) allows data subjects to specify the retention period, which can range from 6 months to 2 years.	
IDIADA	The test dataset includes a recording of 5 minutes CAN data, about 320,000 lines (FSDE), and an anonymized 30min video (FHG).	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	The dataset covers the period from Month 1 (M01) to Month 36 (M36) of the TANGO project.	
MET	2 years historical data will be provided.	
NORB	Yes, the dataset will include personal data and special categories of personal data.	
QBE	**T4.2**	From the duration of the evaluation of the pilots. The data are temporary stored only during the verification process, once the verification is completed, the data are deleted.
	T4.3	The real time data are stored for the duration of the project.
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets from RIAS still under definition (present status to be updated further during the project).	
UMU	Throughout the project, although individual documents may have shorter “expiration”.	
UoG	PEC requests that data subjects specify the scenarios in their use case without including any specific timeframe.	
UoM	Until the end of the project (August 2025).	
UPRC	During the whole period of the project.	
UTH	M24 – M36 of the project, during the pilot evaluation.	
VISAR	Both new environment data and old historical data will be used, and the data will be deleted post the trial.	
VTT	**T3.2**	Not applicable.
	T4.1	Not applicable: data is passport data, and not time bound as such.

1.2. Please indicate the time period covered by the dataset.	
XLAB	Not applicable.

1.3. Does the dataset include personal data (i.e., data that relates to an identified or identifiable individual)? If so, please specify. In your answer, please also specify if the dataset includes "special categories" of personal data (e.g., data revealing racial or ethnic origin, political views, religious beliefs, membership of a trade union). Please also indicate if data concerning children or vulnerable adults will be processed.		
ABI LAB & NTT	Yes, sensitive data includes information about the sender and receiver of the payments, as well as details about the institution.	
ACV	The data collected by the UI component includes the following personal information first name, last name, email, and passport ID.	
AHOP	Yes. Information from hotel guests and passports.	
ANYS	Yes. Information from hotel guests and passports.	
ATOS	**ID**	The data in an ePassport VC includes the ePassport travel document data fields, but it does not include the photograph. Employee information is also included in the EmployeeCredential VC and customer information is included in the CustomerCredential VC. All data to be issued in the ePassport credential is received from the onboarding process which is controlled by the T4.2 ePaspport app. All data to be issued in the EmployeeCredential and CustomerCredential VCs is controlled by the Dashboard app on behalf of the pilot organizations.
	SP	Any data may be encrypted and is pilot dependent (see details from pilot partners' questionnaires regarding what data will be encrypted and possibly on partners encrypting and decrypting such data).
	MLOps	No. Stores mainly metadata.
	AutoML	No. Although it depends on the input data provided by the user (data scientist). The component does not store the data, but rather the resulting trained ML model and metrics. No personal expected.
	FL	No. Although it depends on the input data provided by the user (data scientist). The component does not store the data, but rather the resulting trained ML model and metrics. No personal data expected.
CEA	No.	
CESGA	Yes. Information from hotel guests and passports.	
DBC	Not applicable.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EDAE is not analysing any personal data or information from any of the pilot use cases that participates. No data concerning children or vulnerable adults are processed.	
FHG	Yes.	

1.3. Does the dataset include personal data (i.e., data that relates to an identified or identifiable individual)? If so, please specify. In your answer, please also specify if the dataset includes "special categories" of personal data (e.g., data revealing racial or ethnic origin, political views, religious beliefs, membership of a trade union). Please also indicate if data concerning children or vulnerable adults will be processed.		
FMAKE	No personal data.	
FN	Not applicable.	
FSDE	The test dataset of the CAN data does not include personal data.	
FUJ_LU	The data collected by PAT includes personal information such as first name, last name, email, credit card number, phone number, and passport ID.	
IDIADA	Yes. ePassport data, data from cars.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	The dataset includes limited personal data, such as names and email addresses of newsletter subscribers and event participants. It does not include special categories of personal data. It does not include data concerning children or vulnerable adults.	
MET	Yes, it may contain business or geographic information about METRO customers, but not personal data from those customers.	
NORB	The data present in the ecosystem is encrypted and masked to prevent leakage of PII information and sensitive data.	
QBE	**T4.2**	Yes, the data include the identity information of the person and facial biometrics are extracted. The data are temporary stored only during the verification process, once the verification is completed, the data are deleted. Data concerning children or vulnerable adults will not be processed.
	T4.3	Yes, biometric and behavioural data are processed, which are anonymized where possible; in addition, a unique identifier is used to create the profile, not the actual identity data (data are mapped to a GUID, while name, address, email etc. are not collected). Data concerning children or vulnerable adults will not be processed.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS will still be under definition, although SQD must inform that collection of data such as racial/ethnical/political/religion/association with unions, and others belonging to individuals, are specifically forbidden under Portuguese Law; therefore those information classes are not inquired or collected, and will not be part of any TANGO dataset.	
UMU	**ID**	No, data can be synthetically generated for project usage, although in real world application it could include personal data (depending on requirements of the scenario).
	PEP	No.
	SP	No.
UoG	The data collected by PEC does not includes personal data.	

1.3. Does the dataset include personal data (i.e., data that relates to an identified or identifiable individual)? If so, please specify. In your answer, please also specify if the dataset includes "special categories" of personal data (e.g., data revealing racial or ethnic origin, political views, religious beliefs, membership of a trade union). Please also indicate if data concerning children or vulnerable adults will be processed.		
UoM	No. The data that will be used in our tests with the EDAE will be synthetic data.	
UPRC	No.	
UTH	Yes, it includes data that are used to identify the location of a device, and thus of a person using this device. It does not include any other personal data.	
VISAR	The dataset used in the pilot activities is based on personal data but is, in fact, dummy data. The dummy data is designed to resemble personal data.	
VTT	**T3.2**	Not applicable.
	T4.1	The dataset contains personal data, but no special categories or data of children or vulnerable adults.
XLAB	Not applicable.	

1.4. Are personal data in the dataset pseudonymised? If not, please explain why this is the case.	
ABI LAB & NTT	Yes, sensitive data such as sender and receiver information are pseudonymised , as well as details about the institution.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	No. Data used from the NADIA application supporting the hotel activities.
ATOS	Not applicable.
CEA	Not applicable.
CESGA	No. Real data from customers during the evaluation of the pilot.
DBC	Not applicable.
DUT	No.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Not applicable.
FMAKE	Not applicable.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	The collected data is neither anonymized NORB pseudonymised. The personal data gathered for the TANGO project consists of dummy data used solely for

1.4. Are personal data in the dataset pseudonymised? If not, please explain why this is the case.		
	testing and demonstrating the PAT tool. However, the data is encrypted in the database.	
IDIADA	No. Real data from the evaluation. Anonymization in videos.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Personal data are pseudonymised where feasible, especially in analytics and reporting outputs. Aggregated data do not contain personal identifiers.	
MET	Not applicable.	
NORB	The data present in the ecosystem is encrypted and masked to prevent leakage of PII information and sensitive data.	
QBE	Yes.	
SQD	Not applicable— – SQD will not use its own proprietary datasets, Datasets from RIAS will still be under definition, but personal data collection and sharing will respect Portuguese government issued legislation regarding GPDR principles https://dre.pt/dre/en/detail/act/58-2019-123815982	
UMU	Not applicable.	
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	No, data are used to identify the behaviour and location of users, so cannot be .	
VISAR	Yes.	
VTT	**T3.2**	Not applicable.
	T4.1	No. The use case for the data is identification, so personalized data is necessary.
XLAB	Not applicable.	

1.5. Are data in your dataset anonymised? If not, please explain why you need to use personal data (in other words, why you cannot reach the research objectives by using anonymized data).		
ABI LAB & NTT	It is expected that identifier data will be anonymised. However, since the data is stored within bank environments, anonymization may not be necessary.	
ACV	The collected data is not anonymized.	
AHOP	Yes.	
ANYS	Yes.	
ATOS	**ID**	No, as identity verification requires identifiable attributes. This is a task 4.2 decision and subsequently T4.1 upon analysis of the pilot use cases.

1.5. Are data in your dataset anonymised? If not, please explain why you need to use personal data (in other words, why you cannot reach the research objectives by using anonymized data).		
	SP	This is pilot dependent.
	MLOps	Not applicable.
	AutoML	Not applicable.
	FL	Not applicable.
CEA	Not applicable.	
CESGA	Yes, they have been fully anonymized.	
DBC	Not applicable.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	Personal data which may occur in the original data have been anonymized by IDIADA before sharing the video of the ride. Faces and license plates are blurred, using standard industry methods to comply with the GDPR.	
FMAKE	Not applicable.	
FN	Not applicable.	
FSDE	Personal data, which may occur in the original data have been anonymized by IDIADA before sharing the video of the ride.	
FUJ_LU	The collected data is neither anonymized NORB pseudonymized. The personal data gathered for the Tango project consists of dummy data used solely for testing and demonstrating the PAT tool. However, the data is encrypted in the database.	
IDIADA	Videos anonymised by IDIADA.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Data in public reports and analysis are anonymized or aggregated. Personal data are retained only where necessary for communication purposes, based on consent.	
MET	Yes.	
NORB	The data is anonymized using industry best practice to anonymize PII information and sensitive information. There is a layered and robust data access management system.	
QBE	Yes.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS will still be under definition, but personal data collection and sharing will respect Portuguese government issued legislation regarding GPDR principles https://dre.pt/dre/en/detail/act/58-2019-123815982	
UMU	**ID**	No, as identity verification requires identifiable attributes. Selective disclosure mechanisms (ZKP) limit exposure.

1.5. Are data in your dataset anonymised? If not, please explain why you need to use personal data (in other words, why you cannot reach the research objectives by using anonymized data).		
	PEP	No need to use personal data directly.
	SP	Same as above.
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	No, data are used to identify the behaviour and location of users, so cannot be anonymised.	
VISAR	Yes.	
VTT	**T3.2**	Not applicable.
	T4.1	No. The use case for the data is identification, so personalized data is necessary.
XLAB	Not applicable.	

1.6. Please identify the origin/provenance of the data (how, where, by whom, and when the data was/will be collected, produced, or obtained). Please include in your answer the methods/modes of data collection (e.g., the instruments, hardware, and software used if new data is collected or produced) and explain how data provenance will be documented.		
ABI LAB & NTT	The data originates from banking transactions collected and provided by two collaborating banks. Documentation of data provenance will occur through tracking of data flows and access methods by the participating institutions.	
ACV	No input provided.	
AHOP	Through internet searches and personal interviews with the Experiences providers.	
ANYS	Data from the NADIA platform is collected through sensors that send it to Gateways and send to servers having it stored on ANYS's premises.	
ATOS	**ID**	Pilot issuer applications (ePassport app) and organizations (Tango Dashboard) provide the ePassport and organization member (EmployeeCredential and/or CustomerCredential) information respectively to an SSI Agent (Kotlin application running on a VM) for issuing Verifiable Credentials to persons´ Self-Sovereign Identity Wallets, that are participating in the pilots.
	SP	Data to be encrypted will be generated in the project according to pilot needs. The ABE toolset component creates the ABE public private keys for encrypting and decrypting the data upon request for the sticky policies. The ABE public private keys are stored in the ABE toolset component and will be a few KBs. Public keys are made publicly available for encrypting the data. Electronic files are sent to the ABE toolset component to be encrypted with their associated sticky policies, machine-readable format JSON. The encrypted documents are returned to the publisher and not stored by the ABE toolset component. The reverse process occurs on the ABE toolset component

1.6. Please identify the origin/provenance of the data (how, where, by whom, and when the data was/will be collected, produced, or obtained). Please include in your answer the methods/modes of data collection (e.g., the instruments, hardware, and software used if new data is collected or produced) and explain how data provenance will be documented.		
		to decrypt encrypted files by a consumer providing the encrypted file and the decryption key.
	MLOps	Pilots' data. In the case of the banking pilot, anonymized data (see NTT/ABI answers). In the case of the FMAKE pilot, see FMAKE answers.
	AutoML	Same as above.
	FL	Same as above.
CEA	Data collected in the lab, using the following experimental setup for both datasets: a STM32 Nucleo-144 development board featuring a STM32F756ZG with an ARM Cortex-M7 core is used. Electromagnetic emission is captured, using a Langer RF-B 3-2 electromagnetic probe, a Langer preamplifier PA 303, and a picoscope 5244B. The sampling resolution is set to 8bits, and the sampling frequency is set to 500Msamples/s. The STM32F7 CPU frequency is set to 166.666MHz to capture exactly 3 samples per clock cycle to ease traces analysis. A trigger is placed right before the call to the AES polymorphic instance to ease synchronization.	
CESGA	The data related to customer satisfaction have two sources of origin; during the customer's stay they are collected through written forms and after the stay they are sent an email to fill in the satisfaction questionnaire, this time online.	
DBC	Not applicable.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	**Smart hospitality**	Manual download from the NadiaPlatform developed by UMU.
	Smart manufacturing	Manual download, available online: https://www.makebench.eu/download
	Retail	Provided by METRO end user partner through bilateral communication.
FHG	IDIADA recoded a 30min track of car data during a ride. Exact time and date are not available to Fujitsu. CAN data of 293s of this track are used for the Privacy Risk Score.	
FMAKE	Data captured using PC during machine operations by a machine is located in Leuven, Belgium, operated by an employee of Flanders Make a couple of years ago.	
FN	Not applicable.	
FSDE	IDIADA recorded a 30min track of car data during a ride. Exact time and date are not available to FSDE. CAN data of 293s of this track are used for calculating the attributes for evaluation by TSM.	
FUJ_LU	The data is collected from the data subject through the PAT tool by hotel staff during the registration process or the data subject itself in the smart hospitality	

1.6. Please identify the origin/provenance of the data (how, where, by whom, and when the data was/will be collected, produced, or obtained). Please include in your answer the methods/modes of data collection (e.g., the instruments, hardware, and software used if new data is collected or produced) and explain how data provenance will be documented.		
	use case. This typically occurs at the hotel's front desk or via an online check-in system.	
IDIADA	IDIADA recoded a 30min track of car data during a ride. Exact time and date are not available to Fujitsu. CAN data of 293s of this track are used for the TSM.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Lisbon Council collects and generates the data through dissemination activities. Participant data, newsletter sign-ups, and social media metrics are collected via LIC's website forms, event platforms, and social media analytics tools. Data provenance is documented internally.	
MET	Original data source of transactional data is a custom -made POS system installed in stores and METRO's ERP System for master data. Whenever a customer buys something at a store, full details of the transaction are stored at stores back office and at the end of the day all transactions are transferred to METRO Data warehouse and ERO with a custom ETL procedure. More details (for example, how these data will be available for tango technical partners will be defined).	
NORB	The origin of the data comes from the end user filling in the required information through a customer form for availing a service. The data can also be shared through our data sharing module by integrating to the existing environment through API.	
QBE	For the training of the models, data collected internally at QBE have been used. During the pilots, the users will use the technologies (T4.2 and T4.3) through certain apps at the mobile or web interfaces, through which the data will be collected.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS still under definition (present status to be updated further during the project).	
UMU	**ID**	It will be generated in the project, following pilot needs on which attributes are necessary in their scenario, and generating values through software (randomized) or manual input.
	PEP	It will be generated in the project according to pilot (and general TANGO solution) needs.
	SP	Same as above.
UoG	The data is collected from the user (e.g., engineer, risk manager) through the PEC tool when the user seeks to calculate the privacy risk level of their organisation.	
UoM	Not applicable.	
UPRC	During the first tasks of UPRC, the data were collected by the pilots in the period between November 2022—May 2023 based on case studies, questionnaires and semi-open interviews. During the WP7 implementation and evaluation process, the data will be collected during the period of May to August 2025.	
UTH	Data will be generated through experimentation with devices while designing and prototyping.	

1.6. Please identify the origin/provenance of the data (how, where, by whom, and when the data was/will be collected, produced, or obtained). Please include in your answer the methods/modes of data collection (e.g., the instruments, hardware, and software used if new data is collected or produced) and explain how data provenance will be documented.		
VISAR	The data is entered manually into the platform by the respective user. The data is the identification documents uploaded to the platform by the respective users in digital form, mostly in PDF files, before the visa process. The data is entered into the system by the user, evaluated, and processed. Depending on the respective step, data is either used for filling out the corresponding documents or sent to the relevant stakeholders.	
VTT	**T3.2**	Not applicable.
	T4.1	The data is collected during piloting from users. The data is collected from passports using an android device and application.
XLAB	Not applicable.	

1.7. Will you reuse existing data (including, but not limited to, publicly available data)? If so, specify the source of the data, providing information about the owner/provider of the dataset. Please also explain how you will reuse them and if there are any constraints on the reuse of such data.		
ABI LAB & NTT	No reuse of external or publicly available data is planned.	
ACV	No.	
AHOP	Data from experiences providers are use form training AI algorithm. They will also be reused to personalize tourist experiences.	
ANYS	No.	
ATOS	**ID**	ePassport information is existing data. Organization members in pilots may use existing data also—e.g., email address.
	SP	Unknown but possible, this decision is pilot dependent.
	MLOps	Not applicable; although to test the components available public datasets (e.g., Kaggle examples) have been used.
	AutoML	Same as above.
	FL	Same as above.
CEA	No.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	The encryption/decryption is disposable. No data will be reused. Public data is not involved in our component.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS is not intending to reuse those data for other purposes beyond the TANGO closure.	
FHG	No.	

1.7. Will you reuse existing data (including, but not limited to, publicly available data)? If so, specify the source of the data, providing information about the owner/provider of the dataset. Please also explain how you will reuse them and if there are any constraints on the reuse of such data.		
FMAKE	Yes, data from a previous project will be reused. This data has already been published and it publicly available. There are not limited on the reuse of the data.	
FN	Not applicable.	
FSDE	No.	
FUJ_LU	No.	
IDIADA	No.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Lisbon Council reuses publicly available materials such as templates, reports, and branding elements, ensuring compliance with licensing conditions. Sources are credited.	
MET	Since 2 years historical data will be used in project Tango, these data are already stored in METRO's data warehouse.	
NORB	The information that is already present in government entities and authorities can be reused, the data present if validated can be made available in the ecosystem through having an API interface to reading the dataset.	
QBE	No.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS will be used by SQD during the TANGO project but they are still under definition (present status to be updated further during the project).	
UMU	No.	
UoG	No.	
UoM	No reuse of any data from other partners, the technology will be tested with synthetic data. The only data that may be reused are the accelerated algorithms, which are implemented by UoM and will be open source.	
UPRC	No.	
UTH	No.	
VISAR	No.	
VTT	**T3.2**	Bonner Book Club dataset 2012, for training and evaluating the recommender (hereinafter 'Legacy dataset'); no personal data.
	T4.1	No reuse of an existing dataset.
XLAB	Not applicable.	

1.8. If you have considered but discarded the reuse of existing data, please explain the reasons for that.	
ABI LAB & NTT	No reuse of external or publicly available data is planned.
ACV	Not applicable.

1.8. If you have considered but discarded the reuse of existing data, please explain the reasons for that.	
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Not applicable.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	EXUS will utilize the developed algorithms as a trade secret in order to further expand its product offering called EFS platform. The data refer to different sectors from the ones that EXUS is interested and thus there isn't an intention on re-using those data.
FHG	Not applicable.
FMAKE	No.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	Not applicable.
IDIADA	Not applicable.
INTRA	Not applicable.
KUL	Not applicable.
LC	Reuse of existing materials is considered where relevant.
MET	Not applicable.
NORB	This is not applicable. We are not considering reusing of the information present.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS will be used by SQD during the TANGO project but they are still under definition (present status to be updated further during the project).
UMU	There is no need to reuse data, as these datasets are technical constructs needed to apply the projects' identity management and privacy solutions.
UoG	Because the privacy risk assessment is unique to each user and scenario, there is no need to reuse the data.
UoM	Not applicable.
UPRC	Not applicable.
UTH	No dataset with the requirements sought has been found.

1.8. If you have considered but discarded the reuse of existing data, please explain the reasons for that.		
VISAR	Not applicable.	
VTT	T3.2	Not applicable.
	T4.1	Please provide your answer.
XLAB	Not applicable.	

1.9. To the extent that your dataset contains personal data, has consent been obtained for data collection or is it based on other applicable legal grounds? Have you ensured compliance with the relevant legal principles? Are data subjects in control over their data?		
ABI LAB & NTT	Datasets are in charge of banks.	
ACV	Yes, the use of these data is part of the consent for the onboarding/registration process.	
AHOP	Not applicable.	
ANYS	Yes.	
ATOS	**ID**	The SSI Agent, deployed by pilot partners in their respective premises, will process personal data which has been obtained in a user centric and driven process, but this is outside of the scope of ATOS SSI Agent acting as an issuer. In any case, ATOS signs the DPA for the pilots where ePassport data is required.
	SP	Not applicable.
	MLOps	Same as above.
	AutoML	Same as above.
	FL	Same as above.
CEA	Not applicable.	
CESGA	The only personal data collected by garden hotels is with the express consent of the persons who fill in the questionnaires.	
DBC	Not applicable.	
DUT	Users choose to upload their data in order to use our service. Upon uploading the data, they realize that the data is being processed by TANGO central server. However, the data can be deleted anytime upon users' request. In any case, DUT will sign the DPA for the pilots where the component is required to manage personal data.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Since no personal data were provided to us, there was no need for consent forms.	
FHG	The CAN data for preparing the use case and the scores sent to the TSM do not include any personal data.	
FMAKE	Not applicable.	
FN	Not applicable.	

1.9. To the extent that your dataset contains personal data, has consent been obtained for data collection or is it based on other applicable legal grounds? Have you ensured compliance with the relevant legal principles? Are data subjects in control over their data?		
FSDE	The CAN data for preparing the use case and the scores sent to the TSM do not include any personal data.	
FUJ_LU	The uniqueness of PAT lies in its consent management for collected data. Data subjects have control over their consent and their data. PAT provides a technical solution in compliance with Article 6 of the GDPR, detailing the purpose of the collected data, the data retention period, and the entities with whom the data is shared.	
IDIADA	The CAN data for preparing the use case and the scores sent to the TSM do not include any personal data.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Personal data are processed in compliance with GDPR. Consent is obtained for newsletter subscriptions and event registrations. Data subjects control their data and may withdraw consent.	
MET	No input provided.	
NORB	Data sharing consent management is natively handled by the platform, and integration with existing consent management mechanisms is available through APIs.	
QBE	The legal basis is explicit consent. Data subjects are in control of their data.	
SQD	N/A—SQD will not use its own proprietary datasets. Datasets from RIAS still under definition but personal data collection and sharing will respect Portuguese government issued legislation regarding GPDR principles https://dre.pt/dre/en/detail/act/58-2019-123815982.	
UMU	Not applicable.	
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	Yes.	
VISAR	Data sharing consent management is natively handled by the Fides platform, and integration with existing consent management mechanisms is available through APIs.	
VTT	**T3.2**	Not applicable.
	T4.1	Consent is obtained from users as they are entering their passport data. Data subjects are in control over their data.
XLAB	Not applicable.	

1.10. What is the purpose of the data generation or reuse (in other words, what is the justification for the collection/processing/generation of the data, how and for which purpose will the data be collected/processed/generated/used)? How does it relate to the objectives of TANGO (i.e., for which WP/task/deliverable is it relevant)?		
ABI LAB & NTT	The data is noted to experiment the Federated Learning approach for the fraud detection within the TANGO project, aiming to develop distributed and secure machine learning models that improve fraud detection.	
ACV	Not applicable.	
AHOP	Not applicable.	
ANYS	The data from the NADIA platform is used to gather information about the room status, providing this information to the tourist and the hotel management regarding the use of the rooms by the guests, time spent in the rooms, best cleaning schedules, electricity consumption, etc.	
ATOS	**ID**	It is needed to create and issue Verifiable Credentials to fulfil the project's SSI Task. It is thus relevant in WP4 (specifically T4.1).
	SP	The data generated is needed for fulfilling the project´s ABE Task, specifically WP3 (T3.3).
	MLOps	The metadata generated is needed to keep track of the experiments to generate ML models for explainability and reproducibility. This is in line with WP5 objectives (Obj. 4).
	AutoML	The purpose is to help data scientists to shorten the initial steps on the development of AI models. This is in line with WP5 objectives (Obj. 4).
	FL	The component generates the weights, and the ML model trained in separate datasets. This is in line with WP5 objectives (Obj. 4).
CEA	The data was generated as part of TANGO WP4/T4.5 and aims at assessing whether the considered countermeasures provide security against side-channel attacks.	
CESGA	Customer satisfaction data helps hotel management understand which areas need more attention and improvement in order to provide better service to guests.	
DBC	Not applicable.	
DUT	Encrypted data are generated from the plain data uploaded by users. The purpose is to protect the data from being used by unauthorised users. This process is totally requested by users. This task is related to Deliverable 3.1, 3.2, D2.4, task 3.2 CP-ABE, use case FMAKE-smart manufacturing, WP6 Tango framework, and user-interface.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	The sole purpose of the dataset is showcasing the TSM applied to the TANGO Autonomous Vehicle use case of IDIADA.	
FMAKE	The data has been generated to train and evaluate engineering models that have been created (technology from TANGO-WP5). The data is being used to simulate a real scenario (technology from TANGO-WP3/WP4).	
FN	Not applicable.	
FSDE	The sole purpose of the dataset is showcasing the TSM applied to the TANGO Autonomous Vehicle use case of IDIADA.	

1.10. What is the purpose of the data generation or reuse (in other words, what is the justification for the collection/processing/generation of the data, how and for which purpose will the data be collected/processed/generated/used)? How does it relate to the objectives of TANGO (i.e., for which WP/task/deliverable is it relevant)?		
FUJ_LU	Not applicable.	
IDIADA	The sole purpose of the dataset is showcasing the TSM applied to the TANGO Autonomous Vehicle use case of IDIADA.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	The data support WP8 objectives, disseminating TANGO's findings, promoting engagement, supporting exploitation activities, and contributing to policy recommendations.	
MET	No input provided.	
NORB	The data that will be collected is used for providing the core services, as required by relevant pilots. The details will be used in the WP2 and deliverable D2.1, WP3, and deliverable D3.1, WP7, Task T7.2, and T7.5—Pilots for hospitality and banking.	
QBE	**T4.2**	To perform user identity verification remotely.
	T4.3	To perform user authentication.
SQD	N/A—SQD will not use its own proprietary datasets. Datasets from RIAS will be used by SQD during the TANGO project but they are still under definition (present status to be updated further during the project).	
UMU	**ID**	The data generated is needed to test, including a byproduct of using TANGO's identity management solution: i.e., it is necessary to generate credentials/presentations to fulfil the project's SSI needs. It is thus relevant in WP4 (specifically T4.1) and Obj. 1, Obj. 3. Used indirectly in WP3, T3.3.
	PEP	The data generated is needed for defining the privacy by design solution, and a byproduct of using it (sticky policies and policy enforcement point will let users control access to their data). It is relevant in WP3 (T3.3) and Obj. 1, Obj. 3
	SP	Same as above.
UoG	PEC needs this data to be able to calculate the privacy risk level and recommendations for the TANGO user.	
UoM	The data generated from TornadoVM in T5.3 is to run faster on GPUs and other accelerators. The goal is to run fast algorithms on hardware accelerators (GPUs, CPUs) and measure the energy consumption of those algorithms in order to find the most energy efficient per workload.	
UPRC	The data are related to WP2 (T2.3) and WP7. The purpose of the collected data is to identify the user requirements, build the use cases scenarios, and test the TANGO platform with all pilots.	
UTH	To increase the level of security both for the user and the device, as well as to monitor access to restricted areas.	
VISAR	Not applicable.	
VTT	**T3.2**	VTT internal development of the distributed recommender; validating the results against the legacy measurements of centralized implementation.

1.10. What is the purpose of the data generation or reuse (in other words, what is the justification for the collection/processing/generation of the data, how and for which purpose will the data be collected/processed/generated/used)? How does it relate to the objectives of TANGO (i.e., for which WP/task/deliverable is it relevant)?		
	T4.1	The data is generated and stored to enable secure authentication in TANGO systems. It is relevant especially for WP4 (which is developing the SSI technology) and WP6 (as authentication is a part of the integrated TANGO system).
XLAB	Not applicable.	

1.11. Please explain whether any manipulations to the data(set) will take place during the project (e.g., modifications after original creation, addition of newly collected data, anonymization, etc.).		
ABI LAB & NTT	Expected anonymization.	
ACV	Not applicable.	
AHOP	No.	
ANYS	The dataset is provided raw.	
ATOS	**ID**	Identity data is obtained to create Verifiable Credentials with its own data model and meta data where the personal data is included within a specific "credentialSubject" object. The SSI Agent does not change the personal data but puts it into an electronically signed Verifiable Credential.
	SP	Data is encrypted by ABE toolset but not changed for any of its contents.
	ML Ops	Not applicable. The (meta)data and resulting ML models is generated by the pilots within the project.
	AutoML	Same as above.
	FL	Same as above.
CEA	None.	
CESGA	The dataset is provided raw. Only data from surveys are anonymized.	
DBC	Not applicable.	
DUT	No.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EDAE will conduct only necessary manipulation needed for ensuring the quality of the dataset—e.g., missing values, etc., in order to facilitate the report generation.	
FHG	The TSM and Privacy Risk module do not manipulate the data as it uses only values derived from the dataset, which shall be provided by IDIADA.	
FMAKE	No manipulations are expected. The data will be used as it is.	

1.11. Please explain whether any manipulations to the data(set) will take place during the project (e.g., modifications after original creation, addition of newly collected data, anonymization, etc.).		
FN	Not applicable.	
FSDE	The TSM does not manipulate the data as it uses only values derived from the dataset, which shall be provided by IDIADA.	
FUJ_LU	The collected data is encrypted.	
IDIADA	The TSM does not manipulate the data as it uses only values derived from the dataset, which shall be provided by IDIADA.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Data are updated and modified as dissemination materials evolve. Anonymization and aggregation are applied to engagement metrics.	
MET	No input provided.	
NORB	The data that will be collected will be encrypted in multiple layers to ensure privacy of the data. The technology uses derivation keys to ensure the data is only decrypted at the end point of the browser. Multiple sources are used to ensure randomness of the encryption keys.	
QBE	Anonymization is taking place.	
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets from RIAS are still under definition, RIAS will fully control access as well as any other aspects like transformation, anonymization, storing, and deleting the dataset (present status to be updated further during the project).	
UMU	Datasets will be built and possibly throughout the project as the related technical solutions evolve, and they are applied to pilots.	
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	No.	
VISAR	No.	
VTT	**T3.2**	No changes.
	T4.1	The dataset will be collected as users register for the TANGO platform. No manipulation will be carried out after.
XLAB	Not applicable.	

1.12. Please explain who "owns" the dataset and who will have the rights to control access to the dataset.	
ABI LAB & NTT	Bank institutes.

1.12. Please explain who "owns" the dataset and who will have the rights to control access to the dataset.		
ACV	The organisation (hotel) requesting the data manages access control. However, it is the data subject who grants access by specifying who is authorised to access their data within the hotel.	
AHOP	AHOP.	
ANYS	ANYS.	
ATOS	ATOS is the owner of the software components.	
	ID	Persons own their ePassport data. Pilot organizations own Employee Credentials and CustomerCredentials.
	SP	Pilot publisher organizations making use of ABE toolset are the Owners.
	MLOps	The derivative data from the pilots will be owned by the pilot partner which originally owns the dataset. The resulting data in FL and AutoML components will be ML models owned by the pilots.
	AutoML	Same as above.
	FL	Same as above.
CEA	CEA.	
CESGA	CESGA.	
DBC	Not applicable.	
DUT	Users own the data and have full control to them.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	All three end users for the responsible use cases of Smart Hospitality, Smart Manufacturing, and Retailers are the sole owners of the data providing access to us.	
FHG	IDIADA.	
FMAKE	The dataset is public and available to anyone. Flanders Make is the owner of the dataset.	
FN	Not applicable.	
FSDE	IDIADA.	
FUJ_LU	The organization (hotel) requesting the data manages access control. However, it is the data subject who grants access by specifying who is authorized to access their data within the hotel.	
IDIADA	IDIADA.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Lisbon Council owns the dataset and controls access through its communication team.	

1.12. Please explain who "owns" the dataset and who will have the rights to control access to the dataset.		
MET	MET.	
NORB	The dataset as a whole is not owned by any entity. Instead, each data piece is owned by the data subject and the data control and access is present with them.	
QBE	QBE.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS are still under definition, RIAS will fully control access as well as any other aspects like transformation, anonymization, storing, and deleting the dataset (present status to be updated further during the project).	
UMU	UMU.	
UoG	The data subject grants access by specifying who is authorised to access their data within the TANGO platform.	
UoM	UoG	
UPRC	UPRC.	
UTH	RIAS.	
VISAR	VISAR owns the dataset and has the rights to control access. NORB, as VISAR's technical partner, has access for storage and enabling data access.	
VTT	**T3.2**	VTT.
	T4.1	The data is owned by the data subject. It is stored in a database, the exact details of which will depend on the technical setup of the pilot.
XLAB	Not applicable.	

1.13. Please indicate the partner in charge of collecting, storing, and deleting the data.	
ABI LAB & NTT	The datasets are owned by the collaborating banking institutions, which maintain control over data access.
ACV	Data stored in the database of the user interface will be deleted by ACV within the pilots' premises.
AHOP	AHOP.
ANYS	ANYS.
ATOS	ATOS is the provider of technology deployed and used by pilot partners for processing data as previously described for the SSI Agent, ABE toolset, and the 3 AI support components. Therefore, it is the pilot partners that operate it who are in charge as the controller.
CEA	CEA.
CESGA	CESGA.
DBC	Not applicable.
DUT	Task 3.2 of WP3 will be storing/deleting the encrypted data. However, since the data has been encrypted, only users who own the data has access.
ECO	Not applicable.

1.13. Please indicate the partner in charge of collecting, storing, and deleting the data.		
EGI	Not applicable.	
EXUS	EXUS collects and stores needed data for the purposes of our EDAE module showcase. We will delete the entirety of data beyond project closure.	
FHG	IDIADA (CSD), FHG (SD), Fujitsu (SD).	
FMAKE	Due to its size, the data has been stored by Juelich data service.	
FN	Not applicable.	
FSDE	IDIADA (CSD), FHG (SD), Fujitsu (SD).	
FUJ_LU	The partner responsible for testing the smart hospitality use case will handle collecting, storing, and deleting the data.	
IDIADA	IDIADA (CSD), FHG (SD), Fujitsu (SD).	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Lisbon Council is responsible for collecting, storing, and deleting the data.	
MET	MET.	
NORB	Entities participating in the pilots are in charge of collecting the data. NORB is responsible for deleting the data.	
QBE	QBE.	
SQD	Not applicable—SQD will not its use own proprietary datasets. Datasets from RIAS are still under definition, RIAS will fully control access as well as any other aspects like transformation, anonymization, storing, and deleting the dataset (present status to be updated further during the project).	
UMU	UMU.	
UoG	The partner responsible for testing the public administration and autonomous vehicle use cases will handle collecting, storing, and deleting the data.	
UoM	UoM.	
UPRC	Regarding the WP7, UPRC as a WP Leader will be responsible.	
UTH	UTH.	
VISAR	VISAR is in charge of collecting, storing, and deleting the data. NORB, as VISAR's technical partner, has access for storage and enabling data access.	
VTT	**T3.2**	Not applicable.
	T4.1	The credentials are generated together between the components of QBE, ATOS and VTT. Once the pilots' finished the evaluation, the data in the components from VTT will be deleted.
XLAB	Not applicable.	

1.14. Identify the TANGO partners (if any) having access to/using the dataset during the project. Please explain if they will have access to an aggregated version instead of the full version.	
ABI LAB & NTT	None.
ACV	Not applicable.
AHOP	VTT.
ANYS	In the Smart Hospitality pilot, partners who signed the DPA
ATOS	Pilot dependent.
CEA	All partners for the first one, only CEA for the second one.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Partners involved in Task 3.2 have access to the encrypted data. No one can access the original data.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Pilot dependent.
FHG	FHG and FSDE got copies of the 293s CAN data and the anonymized 30min video for supporting IDIADA in the use case. They shall delete their copies after the project.
FMAKE	All partners have access to the full dataset. In the Smart Manufacturing pilot, partners who signed the DPA
FN	Not applicable.
FSDE	FHG and FSDE got copies of the 293s CAN data and the anonymized 30min video for supporting IDIADA in the use case. They shall delete their copies after the review of the project by the European Commission.
FUJ_LU	The TANGO partner, responsible for the smart hospitality use case, will grant the hotel access to and use of the dataset.
IDIADA	FHG and FSDE got copies of the 293s CAN data and the anonymized 30min video for supporting IDIADA in the use case. They shall delete their copies after the project. In the Automotive pilot, partners who signed the DPA
INTRA	Not applicable.
KUL	Not applicable.
LC	No personal data are shared with other TANGO partners. Aggregated dissemination reports are shared internally.
MET	All the partners developing TANGO's framework for retailers. In the Retail pilot, partners who signed the DPA
NORB	The access to the dataset will be based on the partner being part of the Pilots.
QBE	QBE.
RIAS	In the Smart Manufacturing pilot, partners who signed the DPA
SQD	Not applicable—SQD will not its use own proprietary datasets, Datasets from RIAS are still under definition, RIAS allow access to a curated version of the

1.14. Identify the TANGO partners (if any) having access to/using the dataset during the project. Please explain if they will have access to an aggregated version instead of the full version.		
	RIAS Datasets required for TANGO developments (present status to be updated further during the project).	
UMU	Not applicable.	
UoG	The TANGO partners, responsible for the public administration and autonomous vehicle use cases, will have access to the dataset.	
UoM	Any partner can have access to the source code. The deployment will take place in facilities of the University of Manchester.	
UPRC	Regarding the WP7, UPRC as a WP Leader will be responsible.	
UTH	RIAS, as they will be using the device continuous behavioural authentication module.	
VISAR	NORB, as VISAR's technical partner, has access for storage and enabling data access.	
VTT	**T3.2**	None.
	T4.1	Only VTT should have access to the dataset, but details of access might depend on the technical setups of the pilots.
XLAB	Not applicable.	

1.15. Please identify the partner(s) processing the data included in this dataset and describe the activities that will be performed.		
ABI LAB & NTT	NTT Data will provide some components to process the data in order to experiment the federated learning approach.	
ACV	No partner will process the data; only the hotel will have access to it.	
AHOP	Dataset is going to be processed for training AI algorithm to set the preferences of the tourist.	
ANYS	No input provided.	
ATOS	**ID**	Identity data is processed by the SSI Agent implemented by ATOS in T4.1. The partner responsible for its deployment in a pilot is the responsible for its processing.
	SP	Data to be encrypted/decrypted and data (policies) will be processed by ABE toolset component implemented by ATOS. The partner responsible for its deployment in a pilot is responsible for its processing.
	MLOps	Pilots' partners. Particularly NTT/ABI for the banking pilot and FMAKE in their manufacturing pilot. They would use the components to train ML models using the data at their disposal.
	AutoML	Same as above.
	FL	Same as above.
CEA	Data is processed in order to evaluate the security from the measurements: either attacks (CPA and deep learning attack for the first dataset), both statistical tests (non-specific t-tests for the second dataset) are employed.	
CESGA	Not applicable.	

1.15. Please identify the partner(s) processing the data included in this dataset and describe the activities that will be performed.	
DBC	Not applicable.
DUT	Partners involved in Task 3.2 CP-ABE.
ECO	Not applicable.
EGI	Not applicable.
EXUS	The technical manager is aware for the entirety of processing that takes place from the partners.
FHG	IDIADA will provide attribute related scores based on the test data. For privacy risk, scoring IDIADA will rely on the related tool of FHG. FTSDE will use the scores to calculate a weighted Trustworthiness Score.
FMAKE	The data will be used by most technical partners represented in WP3, WP4, and WP5 of TANGO.
FN	Not applicable.
FSDE	IDIADA will provide attribute-related scores based on the test data. For privacy risk, scoring IDIADA will rely on the related tool of FHG. FTSDE will use the scores to calculate a weighted Trustworthiness Score.
FUJ_LU	No partner will process the data; only the hotel will have access to it.
IDIADA	IDIADA will provide attribute-related scores based on the test data. For privacy risk, scoring IDIADA will rely on the related tool of FHG. FTSDE will use the scores to calculate a weighted Trustworthiness Score.
INTRA	Not applicable.
KUL	Not applicable.
LC	LC processes the data. Activities include content creation, event management, analysis of dissemination metrics, anonymisation, and reporting.
MET	EXUS will process and analyse the dataset.
NORB	The entities that are part of the agreement and are part of the pilots will be able to view and perform actions to the datasets based on the privileges.
QBE	QBE.
SQD	Not applicable—SQD will not its use own proprietary datasets. During the TANGO project, SQD will assist the RIAS demonstrator with their Dataset Management, including but limited to preparing specific RIAS-TANGO Project datasets, with regard to EU Data legislation Compliance. This means separating specific TANGO Datasets, adding persistent identifiers, implementing standard metadata access protocols, and, if necessary, transforming data for compliance with FAIR principles (present status to be updated further during the project).
UMU	It is processed by the tools generated in tasks T3.3 and T4.1. Specifically, data (policies) in SP will be processed by ABE toolset by ATOS.
UoG	No partner will process the data.
UoM	UoM.
UPRC	Not applicable.

1.15. Please identify the partner(s) processing the data included in this dataset and describe the activities that will be performed.		
UTH	UTH will use the data that are related to device operational characteristics to provide device continuous authentication.	
VISAR	The specific activities performed include data entry, evaluation, and processing for visa/immigration applications. Depending on the step, data is used for filling out documents or sent to relevant stakeholders.	
VTT	**T3.2**	Only VTT.
	T4.1	After generation, the data will be processed by data space connector components (deployed by INTRA and UMU [PEP/PDP]) and the TANGO Dashboard (UMU).
XLAB	Not applicable.	

1.16. Please explain whether the dataset will be combined with any other datasets (from TANGO partners or third parties).	
ABI LAB & NTT	Not applicable.
ACV	No data will be combined.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	No.
CEA	Not applicable.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Not applicable.
FMAKE	No.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	No data will be combined. If PEP/PDP integration is successful, then there can be a demonstration use case with dummy data to demonstrate the combination of room sensor data with a client's preferences data. This is for UMU to clarify.
IDIADA	Not applicable.
INTRA	Not applicable.
KUL	Not applicable.

1.16. Please explain whether the dataset will be combined with any other datasets (from TANGO partners or third parties).		
LC	The dataset is not combined with other datasets containing personal data. Aggregated metrics may be compared with project KPIs.	
MET	No.	
NORB	The dataset is unique and will not be combined with any other dataset.	
QBE	No.	
SQD	Not applicable—SQD will not use its own proprietary datasets, RIAS curated datasets will be available for necessary project activities.	
UMU	No.	
UoG	Data will not be combined.	
UoM	It will be operated along with EDAE (T5.1).	
UPRC	Not applicable.	
UTH	Not applicable.	
VISAR	Not applicable.	
VTT	**T3.2**	No.
	T4.1	The data will not be combined.
XLAB	Not applicable.	

1.17. To whom may the data be useful outside of the TANGO project ('data utility')?		
ABI LAB & NTT	None.	
ACV	No one.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Identifier data are not shared.
	SP	Public data examples (e.g., schema definitions) that do not contain personal data of the TANGO technology and documenting how it works could be made available.
	MLOps	Not applicable.
	AutoML	Not applicable.
	FL	Not applicable.
CEA	The datasets may be useful to researchers in the side-channel community.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Not applicable.	

1.17. To whom may the data be useful outside of the TANGO project ('data utility')?		
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	The particular datasets were specific to the use cases of Smart Manufacturing, Smart Hospitality, and Retailers and thus it is not expected to be useful to other partners outside TANGO project.	
FHG	No access to the dataset planned for parties outside of the use case.	
FMAKE	Companies or research institutes interested in 3D printing.	
FN	Not applicable.	
FSDE	No access to the dataset planned for parties outside of the use case.	
FUJ_LU	No one, this is outside of the project scope for PAT.	
IDIADA	No access to the dataset planned for parties outside of the use case.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Policymakers, researchers, and organizations interested in digital infrastructures, secure data exchange, and data spaces may find dissemination outputs useful.	
MET	No one outside TANGO will be able to use the data.	
NORB	The PII information that is collected can be used by any utility that may require details for identity verification and processing of personal data.	
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets, RIAS TANGO Datasets will be useful for other industrial manufacturing facilities pursuing solutions for similar problems.	
UMU	It may be useful as examples of the TANGO technology and for documenting how it works.	
UoG	Not applicable.	
UoM	Any enterprises that have Java implementations that can run on hardware accelerators.	
UPRC	Not applicable.	
UTH	To research and academic communities to further conduct research.	
VISAR	Responsible authorities involved in the visa/immigration process.	
VTT	**T3.2**	No requests have been presented.
	T4.1	No one, the data is meant to be used only by data subjects during TANGO pilots.
XLAB	Not applicable.	

1.18. Will personal data included in your dataset be transferred to countries outside the EU/ EEA? If so, please specify.	
ABI LAB & NTT	No.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	No.
CEA	Not applicable.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	No.
FMAKE	No.
FN	Not applicable.
FSDE	No.
FUJ_LU	No.
IDIADA	No.
INTRA	Not applicable.
KUL	Not applicable.
LC	Personal data are not transferred outside the EU/EEA.
MET	No.
NORB	The data will remain within the EU/EEA.
QBE	No.
SQD	Not applicable—SQD will not use its own proprietary datasets, RIAS will have full control of their own datasets; transfer of RIAS Data to countries outside the EU is not foreseen under the TANGO SOW.
UMU	No.
UoG	No.
UoM	Not applicable.
UPRC	No.

1.18. Will personal data included in your dataset be transferred to countries outside the EU/ EEA? If so, please specify.		
UTH	No.	
VISAR	No.	
VTT	**T3.2**	No
	T4.1	No.
XLAB	Not applicable.	

Annex III – TANGO Data Storage, Quality and Sharing between TANGO Partners

2.1. Please explain how and where the data will be stored during the project (e.g., local storage, cloud storage, storage media etc.).		
ABI LAB & NTT	The data will be stored in the bank infrastructure.	
ACV	The data will be stored in the local storage of the Smart Hospitality Pilot.	
AHOP	Experiences dataset is locally stored.	
ANYS	NadiaPlatform stores the data for 7 days in the gateways. Persistence data are stored on ANYS premises.	
ATOS	**ID**	No user data is permanently stored in the SSI Agent, only session data for issuing a VC.
	SP	No pilot supplied data is permanently stored in the ABE toolset component, only session data for encrypting/decrypting data. The ABE public/private keys are stored in the ABE toolset component for the duration of the pilot.
	MLOps	Components storing metadata and ML models trained by Pilots' partners on their own local storage. Particularly NTT/ABI for the banking pilot and FMAKE in their manufacturing pilot.
	AutoML	Same as above.
	FL	Same as above.
CEA	Storage on CEA's servers, in a folder dedicated to the project.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Local storage from users.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Local storage.	
FHG	For analysis of the showcase: local copies at FHG, FTSDE, IDIADA. For the showcase itself: data will be processed and stored on IDIADA premises.	
FMAKE	The data remains available through the public website. Partners may store the data locally in their environment.	
FN	Not applicable.	
FSDE	For analysis of the showcase: local copies at FHG, FTSDE, IDIADA. For the showcase itself: data will be processed and stored on IDIADA premises.	
FUJ_LU	The data will be stored in the local storage of the smart hospitality pilot.	
IDIADA	For analysis of the showcase: local copies at FHG, FTSDE, IDIADA. For the showcase itself: data will be processed and stored on IDIADA premises.	
INTRA	INTRA does not store any data. The data generated during the Pilot demonstrations are handled and stored by the respective Pilot Leaders.	

2.1. Please explain how and where the data will be stored during the project (e.g., local storage, cloud storage, storage media etc.).		
KUL	Not applicable.	
LC	Website or other relevant data are stored on LC's secure local servers and backed up on OwnCloud.	
MET	All original data are stored on premise at METRO local storage.	
NORB	The data will be stored in Fides proprietary distributed storage DRILL, which uses cloud storage as a data backend.	
QBE	**T4.2**	Data are not stored.
	T4.3	Microsoft Azure in Netherlands.
SQD	Not applicable—SQD will not use its own proprietary datasets, datasets to be handled from RIAS will be kept and stored by RIAS (present status to be updated further during the project).	
UMU	During execution, documents will be stored locally by the corresponding actors (e.g., users store their credentials). Examples and demonstrators will be uploaded to project's cloud data tools.	
UoG	The data will be stored on OwnCloud.	
UoM	The synthetic data for testing are stored in local storage.	
UPRC	Local storage will be in the facilities of UPRC and on OwnCloud.	
UTH	Local Storage (in a local server located at RIAS premises).	
VISAR	Data will be stored within the TANGO component Fides, specifically in its distributed legally compliant data backend DRILL, using an internal Fides format.	
VTT	**T3.2**	Local storage.
	T4.1	This depends on the technical setup of the pilots. So far it has been stored on a Kubernetes cluster set up by WP6.
XLAB	Not applicable.	

2.2. Please provide information on backups (e.g., how/where the data will be backed up, how often the backup will be performed, etc.).	
ABI LAB & NTT	Not applicable.
ACV	No backups.
AHOP	Each time a record is updated.
ANYS	Nadia platform persistence data are backup once a day.
ATOS	Not applicable.
CEA	Backups occur every day.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.

2.2. Please provide information on backups (e.g., how/where the data will be backed up, how often the backup will be performed, etc.).		
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS is storing data backup in its internal SharePoint that has multi-factor authentication. EXUS is an ISO:27001 certified company and follows the corresponding policies for high level of security of information.	
FHG	Not applicable.	
FMAKE	The data is public on a data service and Flanders Make keeps a local copy.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	Not applicable.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Automated weekly backups are performed and securely stored.	
MET	Backup of original data is performed on daily schedule.	
NORB	Fides DRILL is a highly distributed and redundant storage, naturally serving as a backup.	
QBE	**T4.2**	Not applicable.
	T4.3	Continuous backup is performed for 7 days.
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS (present status to be updated further during the project).	
UMU	Same backup plan as project's data sharing tools.	
UoG	Aligned with TANGO's backup policy.	
UoM	Not applicable.	
UPRC	The backups will be performed once a week and there will be established in the facilities of UPRC.	
UTH	Not applicable.	
VISAR	Backups are automated right before essential system updates and data schema migrations. Access to backups is restricted only to developers' team members participating in maintaining the particular services.	
VTT	**T3.2**	Local backups
	T4.1	No backups.
XLAB	Not applicable.	

2.3. Please provide information about the measures adopted to ensure data integrity, data quality and confidentiality of data.		
ABI LAB & NTT	Not applicable.	
ACV	Not applicable.	
AHOP	Not applicable.	
ANYS	An etl process guarantees the date quality at the gateways.	
ATOS	**ID**	Data integrity will be assured through the cryptographic signature proofs used for their generation as associated to their Decentralized Identifiers. Data quality aspects are determined by T4.2.
	SP	ABE encryption provided by the ABE toolset component ensures the confidentiality of the data.
	MLOps	Data integrity will be ensured by the way the components handle their data (e.g., MLFlow, MLJar of the FL framework own ways of data handling). The components work with data locally.
	AutoML	Same as above.
	FL	Same as above.
CEA	**Data integrity**: Backups are done as described. **Data quality**: The electromagnetic probe is positioned carefully to provide good data quality. Position quality for hiding countermeasures was checked on unprotected implementations: the position should allow easy exploitation of side-channel leakage on such implementations. Position quality for masking countermeasures was checked on protected implementation by disabling the random number generator: in such context, we should detect high side-channel leakage. **Data confidentiality**: Data is stored on folders with access control to allow access to relevant people only, or it is made public when possible from a confidentiality point of view.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	The data is encrypted to ensure the confidentiality. The hash value is used to verify the integrity of data. The wrong hash value (i.e., the file is revised by malicious users) will result in failed decryption.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS is an ISO:27001 certified company and follows the corresponding policies for high level of security (confidentiality) and data integrity of information and only authorized personnel have access to the data. EXUS's EDAE objective is to ensure a high level of data quality in order to feed the ML components of TANGO.	
FHG	For preparation of the showcase: IDIADA, FHG, and FTSDE ISM and GDPR-related measures apply ensuring data integrity, quality, and confidentiality. During the showcase, the IDIADA ISM and GDPR related measures apply. The TSM module itself is containerized and deployed in the secure TANGO or IDIADA environment.	

2.3. Please provide information about the measures adopted to ensure data integrity, data quality and confidentiality of data.		
FMAKE	Quality checks were performed before the data is released. Hashes were provided as well.	
FN	Not applicable.	
FSDE	For preparation of the showcase: IDIADA, FHG, and FTSDE ISM and GDPR-related measures apply ensuring data integrity, quality, and confidentiality. During the showcase, the IDIADA ISM and GDPR related measures apply. The TSM module itself is containerized and deployed in the secure TANGO or IDIADA environment.	
FUJ_LU	Data Integrity: PAT implements robust validation checks and regular audits to maintain the accuracy and consistency of the data throughout its lifecycle. Data Quality: The data collection processes of PAT are designed to minimize errors and ensure that the data is complete, accurate, and up to date. Confidentiality: PAT employs an encryption technique and access control mechanisms to protect the data from unauthorized access. Only authorized personnel, as specified by the data subjects, can access the data.	
IDIADA	For preparation of the showcase: IDIADA, FHG, and FTSDE ISM and GDPR-related measures apply ensuring data integrity, quality and confidentiality. During the showcase, the IDIADA ISM and GDPR related measures apply. The TSM module itself is containerized and deployed in the secure TANGO or IDIADA environment.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Data integrity is ensured through encryption, access controls, and regular monitoring. Confidentiality is maintained by limiting access to authorized LIC personnel—for the information stored in LC repositories.	
MET	To be discussed.	
NORB	The data will be encrypted.	
QBE	**T4.2**	Data are not stored.
	T4.3	Data are stored on the cloud.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be curated previously to being used in the TANGO Project, with the goals of supplying anonymized, homogenous, datasets to the required TANGO activities (present status to be updated further during the project).	
UMU	Data integrity will be assured through the cryptographic means used for their generation (e.g., p-ABC signatures), documents stored and shared will not contain confidential data.	
UoG	Data Integrity: PEC implements robust validation checks and regular audits to maintain the accuracy and consistency of the data throughout its lifecycle. Data Quality: The data collection processes of PEC are designed to minimise errors and ensure that the data is complete, accurate, and up to date. Confidentiality: PEC employs access control mechanisms to protect the data from unauthorised access. Only authorised personnel, as specified by the data subjects, can access the data.	

2.3. Please provide information about the measures adopted to ensure data integrity, data quality and confidentiality of data.		
UoM	Not applicable.	
UPRC	In order to reassure the data integrity and their quality, researchers will be in constant communication with all the pilots.	
UTH	To ensure integrity, access control is ensured, and users are instructed to follow the guidelines about how to use the tools and services that handle data. Moreover, input data is validated to ensure their accuracy and veracity of information regarding the recorded values. Data confidentiality is achieved through enabling secure encryption to the collected data and allowing access through authorization to users according to their respective roles. Ensuring data quality stems from the combination of the aforementioned processes along with the regular security checks and audit controls.	
VISAR	Backups: Regular backups are performed to maintain data integrity and ensure that data can be restored in case of loss or corruption. Data Validation: Procedures are in place to validate data upon entry to ensure accuracy and quality. Security Protocols: Robust security measures such as encryption, secure access controls, and regular security audits to protect the data against unauthorized access and breaches are implemented. Usage of Fides Tango component, specifically its DRILL data backend, provides all of the above properties.	
VTT	**T3.2**	No personal data included.
	T4.1	Software development best practices.
XLAB	Not applicable.	

2.4. Please provide an overview of the measures that you have or will put in place to ensure data security (including data recovery, as well as secure storage/archiving and transfer of sensitive data). Please include in your answers information about, but not necessarily limited to, physical security, network security, and security of computer systems and files.		
ABI LAB & NTT	Not applicable.	
ACV	Data is only accessible through tokens provided by a security application.	
AHOP	Not applicable.	
ANYS	Data is only accessible through tokens provided by a security application.	
ATOS	**ID**	SSI Agent issues all credentials with cryptographic proofs to ensure the data's integrity and root of trust by a trusted issuer. The trusted issuer´s (SSI Agent) Decentralized Identity will be added to the Data Space Connector´s Trusted Issuer List. Additionally, HTTPS is used to ensure that REST API calls to the SSI Agent are encrypted.
	SP	ABE encryption provided by the ABE toolset component ensures the confidentiality of the data. Additionally, HTTPS is used to ensure that REST API calls to the ABE toolset are encrypted.
	MLOps	Deployed locally within the pilot's organisations or even within data scientist equipment.
	AutoML	Same as above.

2.4. Please provide an overview of the measures that you have or will put in place to ensure data security (including data recovery, as well as secure storage/archiving and transfer of sensitive data). Please include in your answers information about, but not necessarily limited to, physical security, network security, and security of computer systems and files.		
	FL	Data from pilots is only accessed locally. Only ML models weights are shared among the participants of the federated learning network. Modules for privacy available within the tool if needed.
CEA	Side-channel measurements are not sensitive data, and the results, although valuable from a scientific point of view, do not put at risk any entity or person. However, the following measures are in place. CEA is physically accessible only though access control. Computer disks are all encrypted. Sensitive data can be sent though encrypted file transfer. A VPN is used to access data located on CEA-servers. Backups are accessible only by IT services.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	The file will be split into several data chunks, and all chunks are encrypted with AES256 algorithm.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS as an ISO:27001 certified company follows the corresponding policies for high level of security, including, but not limited to: • Firewall and Intrusion Detection/Prevention. • Regular Penetration Testing. • Access Control & Least Privilege Principle: User access is role-based and reviewed regularly. Multi-factor authentication (MFA) is enforced. • Endpoint Protection: Antivirus, anti-malware, and endpoint detection and response (EDR) solutions are deployed on all devices. • Secure File Transfer Protocols: Secure FTP (SFTP) and Virtual Private Networks (VPNs) are used for file transfers. • Security Audits: Regular ISO 27001 internal and external audits assess compliance with security policies. • Security Awareness Training: Employees undergo security training to recognize and prevent cyber threats.	
FHG	The TSM data are local in time. There is no need for persistence as the data may be reproduced easily.	
FMAKE	There is not sensitive data. The data is stored locally in our company network, so security is handled by the IT department. This will also act as data recovery in case the online server fails.	
FN	Not applicable.	
FSDE	The TSM data are local in time. There is no need for persistence as the data may be reproduced easily.	
FUJ_LU	To ensure data security, PAT has implemented comprehensive measures that safeguard data, even during transfers between TANGO components. These measures include robust authentication protocols and encrypted storage solutions.	
IDIADA	The TSM data are local in time. There is no need for persistence as the data may be reproduced easily.	

2.4. Please provide an overview of the measures that you have or will put in place to ensure data security (including data recovery, as well as secure storage/archiving and transfer of sensitive data). Please include in your answers information about, but not necessarily limited to, physical security, network security, and security of computer systems and files.		
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Physical, network, and system security measures are applied, including firewalls, password protection, secure file transfer protocols, and anti-virus software.	
MET	To be discussed.	
NORB	The methodology that is used ensured that the data is encrypted in multiple layers, and decryption to have the complete data displayed is handled through the randomness of the technology.	
QBE	QBE is certified for ISO27001 and ISO22301.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS (present status to be updated further during the project).	
UMU	**ID**	During execution, protected through the project's SSI solution (e.g., encrypted storage). Regarding examples/demonstrators uploaded, they will contain no sensitive data and benefit from the security measures of the project's data sharing tool.
	PEP	During execution, policies will be fully public and will rely on the project's data sharing tools for transit. Regarding examples/demonstrators uploaded, they will contain no sensitive data and benefit from the security measures of the project's data sharing tool.
	SP	Same as above.
UoG	To ensure data security, PEC implements appropriate authentication measures and utilises the TANGO data recovery plans and policies.	
UoM	No processing of sensitive data. However, the access to local server is protected by VPN and access is only possible through encrypted keys (SSH).	
UPRC	There are guards all day long in the facilities and the entrance is forbidden to anyone with no access. Moreover, the access to local pcs will be with password and the files will be encrypted in order to prevent any security bridge. Last but not least, to reassure that there will be no hacks, the local storage will be in hard drives with no internet access.	
UTH	No personal data or confidential information will be disclosed due to data anonymization (dummy data). This measure effectively mitigates the risk of unauthorized disclosure while facilitating necessary data sharing for the project's purposes.	
VISAR	**T3.2**	No personal data included
	T4.1	Standard software development best practices.
VTT	Not applicable.	
XLAB	Not applicable.	

Annex IV – Data Availability and Sharing between TANGO Partners

3.1. For datasets that are available to other TANGO partners, please provide information about **conditions for or restrictions to access to or use** of the dataset (including, but not limited, to control of access to personal data/confidential information).		
ABI LAB & NTT	Access to data is exclusively provided to the banking institutions and NTT Data, responsible for supplying components for dataset processing and the Federated Learning experiment.	
ACV	No datasets are available to other TANGO partners.	
AHOP	The dataset is sent by email.	
ANYS	All data are provided via API with security token access.	
ATOS	**ID**	Not applicable.
	SP	Pilot dependent.
	MLOps	All partners will have access to the components for deployment. Data will be only available to the specific pilots' own data where the components are deployed.
	AutoML	Same as above.
	FL	Same as above.
CEA	The side-channel datasets are accessible to CEA only, either made public. There is no particular restriction to use them.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Only authorised users from TANGO portal can access SEDSS service and upload/download data. The key for decryption is further protected by CP-ABE. Only authorized users can retrieve the key for decryption.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	The other partners—e.g., SQD for the UI component—are viewing the results of our data processing in other formats—e.g., JSON or Image file reports. They do not have a direct access to the original dataset.	
FHG	No datasets but only the results of the showcase are available to other TANGO partners outside of the IDIADA use case.	
FMAKE	Public dataset is licensed as CC BY-NC-4.0.	
FN	Not applicable.	
FSDE	No datasets but only the results of the showcase are available to other TANGO partners outside of the IDIADA use case.	
FUJ_LU	No datasets are available to other TANGO partners.	
IDIADA	No datasets but only the results of the showcase are available to other TANGO partners outside of the IDIADA use case.	

3.1. For datasets that are available to other TANGO partners, please provide information about **conditions for or restrictions to access to or use** of the dataset (including, but not limited, to control of access to personal data/confidential information).		
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Only anonymised and aggregated dissemination reports are shared internally. Personal data are not shared.	
MET	To be discussed.	
NORB	Partners of the banking and hospitality pilots have access to the data.	
QBE	Not available to other TANGO partners.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be curated previously to being used in the TANGO Project, with the goals of supplying anonymized, homogenous datasets to the required TANGO activities (present status to be updated further during the project)	
UMU	All partners will have access to documents published in the project's sharing tool, which will contain no personal or confidential information.	
UoG	Not available to other TANGO partners.	
UoM	The dataset are open-sourced. They can be used by anyone.	
UPRC	LSTECH and UTH will have access to the data for the activities of WP2. NETCOMPANY has access to the WP7 related information.	
UTH	Data is not intended for use besides UTH and RIAS.	
VISAR	Access to real personal data is restricted.	
VTT	**T3.2**	Not applicable.
	T4.1	Not available to other TANGO partners.

3.2. Please provide technical information on access to data (e.g., tools or software needed to access and use the data), the record-keeping of identities, data alteration, and access events, and the protocols for access to data (e.g., Is authentication needed? Is there a data access request procedure?)		
ABI LAB & NTT	Not applicable.	
ACV	Access to the configuration file of the SQLite database.	
AHOP	Not applicable.	
ANYS	For accessing data authentication is needed by Fiware Keyrock.	
ATOS	**ID**	ePassport app authenticates the ePassport before requesting the SSI Agent to issue an ePassport VC. The SSI Agent verifies the ePassport Verifiable Presentation (by checking its signature against the wallets public key associated to its DID) and requests the users' associated organization data (REST API) to the Dashboard app based on the personal identifier in the ePassport, before issuing an Employee Credential or Customer Credential VC with the users´ organization

3.2. Please provide technical information on access to data (e.g., tools or software needed to access and use the data), the record-keeping of identities, data alteration, and access events, and the protocols for access to data (e.g., Is authentication needed? Is there a data access request procedure?)		
		data. The employee and customer data resides in the dashboard application of the issuer and is outside the control of the SSI Agent.
	SP	Access control is handled by the controller SP component and NOT by the ATOS ABE toolset.
	MLOps	Local access.
	AutoML	Same as above.
	FL	The client nodes are deployed in specific client nodes with access to the training data. The client nodes must be identified and with network communication over http(s) to the central node to average the weights of the trained model. Only the weights are shared from a local node to the central node and vice versa.
CEA	Authentication is needed to access the second dataset. The first one is public. Both datasets use standard formats, making technical use easy, such as the widely used NumPy file format. Data access can be requested by email.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	With respect to access to data, users can keep encrypted data in local environment, which doesn't require extra tools to access the data; or they keep the encrypted data in local server with their own API, then the authentication is required from user side.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	**Smart hospitality**	Nadia platform: https://www.nadiaplatform.com/iot/login.html Requires authentication
	FMAKE	https://www.makebench.eu/download Public dataset
FHG	The regular measures of IDIADA, FHG and Fujitsu apply.	
FMAKE	The data can be downloaded directly from the website using https protocol.	
FN	Not applicable.	
FSDE	The regular measures of IDIADA, FHG, and Fujitsu apply.	
FUJ_LU	Access to data is managed through a combination of specialized tools and software, such as Keycloak for authentication and access control. All access events, including identities of users accessing the data, data alterations, and access events, are meticulously recorded to ensure traceability and accountability. Authentication is mandatory for accessing the data, ensuring that only authorized personnel can gain entry.	
IDIADA	The regular measures of IDIADA, FHG, and Fujitsu apply.	
INTRA	Not applicable.	
KUL	Not applicable.	

3.2. Please provide technical information on access to data (e.g., tools or software needed to access and use the data), the record-keeping of identities, data alteration, and access events, and the protocols for access to data (e.g., Is authentication needed? Is there a data access request procedure?)		
LC	Dissemination reports are shared in standard formats (PDF, CSV). Access is controlled by LC's personnel and project secure repositories.	
MET	Partners developing TANGO's framework for retailers.	
NORB	Access is governed by NORB Fides proprietary DnAMS tool.	
QBE	MFA is required and there is a data request procedure based on the privacy policy.	
SQD	N/A—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS, access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).	
UMU	Same requirements as project's data sharing tool.	
UoG	All access events, including identities of users accessing the data, data alterations, and access events, are meticulously recorded to ensure traceability and accountability. Authentication is mandatory for accessing the data, ensuring that only authorised personnel can gain entry.	
UoM	GitHub access, not applicable.	
UPRC	Not applicable.	
UTH	Authentication and authorization mechanisms of the database system used (Influx DB).	
VISAR	Data will be stored within the TANGO component Fides, specifically in its distributed legally compliant data backend DRILL, using an internal Fides format. The Fides API document (currently a work in progress) will provide information on available APIs/types of queries for data retrieval.	
VTT	**T3.2**	Not applicable.
	T4.1	Not applicable: data is only accessible to data subject.
XLAB	Not applicable.	

3.3. Please explain the mechanisms put in place to ensure that data is available when needed to authorised users for the purposes of TANGO.		
ABI LAB & NTT	Not applicable.	
ACV	No measures foreseen as of now.	
AHOP	Not applicable.	
ANYS	All data are protected.	
ATOS	**ID**	Pilot dependent.
	SP	Same as above.
	MLOps	Available upon deployment in the pilots' premises.
	AutoML	Same as above.

3.3. Please explain the mechanisms put in place to ensure that data is available when needed to authorised users for the purposes of TANGO.		
	FL	Same as above.
CEA	When in need to share particular data to partners, this took place either through OwnCloud.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Total reliance on TANGO authorisation mechanism.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EDAE follows the TANGO architecture for providing available data to other users.	
FHG	No input provided.	
FMAKE	Data is available without authorisation.	
FN	Not applicable.	
FSDE	The extracted data may be shared by IDIADA on request.	
FUJ_LU	To ensure data availability for authorized users within the TANGO project and especially within the smart hospitability pilot (hotel), PAT has several mechanisms. These include robust access control systems, data, and consent management. Authorized users can access the data through secure authentication protocols, ensuring that data is readily available when needed for hotel purposes.	
IDIADA	No input provided.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Reports and materials are made available via secure internal channels and project meetings, ensuring timely access.	
MET	Not yet defined.	
NORB	Access is governed by NORB Fides proprietary DnAMS tool.	
QBE	Data can be requested through dpo@quadible.co.uk or through QBE website www.quadible.co.uk	
SQD	N/A—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).	
UMU	Available through OwnCloud.	
UoG	These include robust access control systems. Authorised users can access the data through secure authentication protocols ensuring that data is readily available when needed.	
UoM	Not applicable.	
UPRC	Available through the project's tools.	

3.3. Please explain the mechanisms put in place to ensure that data is available when needed to authorised users for the purposes of TANGO.		
UTH	Database deployed on a highly available Kubernetes system.	
VISAR	To be determined.	
VTT	**T3.2**	Requests to VTT, followed by sending the Legacy dataset to the requester.
	T4.1	Not applicable: data is only accessible to data subject.
XLAB	Not applicable.	

Annex V – Archiving, Preservation and Deletion of TANGO Data

4.1. Please explain what will happen to the data after the end of the project. In doing so, please set out, among others, how/where the data will be kept (e.g., archive, trusted repository for long-term preservation and curation), for how long the data will be kept, in what format they will be kept, how they will be deleted (e.g., according to which technical standard), etc. If not all data will be kept, please explain which data will be kept and how the selection of data for preservation will take place. Please also explain why some data cannot be preserved (e.g., legal or contractual restrictions, storage/budget issues, institutional policies, etc.). Please also specify in your answers for how long any backups will be held and how they will ultimately be destroyed.		
ABI LAB & NTT	Each institution collects data locally, ensuring confidentiality and security of the information during processing.	
ACV	No data will be retained after the end of the project. Throughout the project, only dummy data will be used.	
AHOP	Data of the pilot deleted after the evaluation.	
ANYS	Data of the pilot deleted after the evaluation.	
ATOS	**ID**	No data needs to be kept after the project ends. The pilot responsible is in charge of ensuring that all data, config. data, session data in DBs, event logs etc. shall be destroyed. Even after deleting files, remnants may remain in the VM's free space. Pilot administrators should use tools to overwrite the free space on the VM's virtual disks.
	SP	Same as above.
	MLOps	The resulting ML models will be proprietary to the project partners, who would license them at their own accord. The metrics and metadata generated during the training phase and persisted in the components (mainly applied to MLOps) will remain at the disposal of the pilots in their own storage. No other data will be kept.
	AutoML	Same as above.
	FL	Same as above.
CEA	The side-channel datasets will be deleted if they have not been shared with the community, after it is clear we won't analyse it again (dataset will thus be useless at that point). The result graphs will have been published and will be kept. CSV files will be used to keep the results.	
CESGA	Data of the pilot deleted after the evaluation.	
DBC	Not applicable.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS will delete any TANGO related data beyond project closure. Data stored in cloud environments is securely deleted per the cloud provider's ISO 27001 and SOC 2-compliant processes. Data deletion activities are logged and audited to verify compliance with security policies. Periodic internal and external audits	

4.1. Please explain what will happen to the data after the end of the project. In doing so, please set out, among others, how/where the data will be kept (e.g., archive, trusted repository for long-term preservation and curation), for how long the data will be kept, in what format they will be kept, how they will be deleted (e.g., according to which technical standard), etc. If not all data will be kept, please explain which data will be kept and how the selection of data for preservation will take place. Please also explain why some data cannot be preserved (e.g., legal or contractual restrictions, storage/budget issues, institutional policies, etc.). Please also specify in your answers for how long any backups will be held and how they will ultimately be destroyed.	
	ensure that data disposal processes align with ISO 27001 and relevant regulatory requirements.
FHG	No data will be retained after the end of the project. Throughout the project, only a test dataset will be used.
FMAKE	It is expected that the data will remain available for 3 years after the end of the project.
FN	Not applicable.
FSDE	No data will be retained after the end of the project. Throughout the project, only a test dataset will be used.
FUJ_LU	No data will be retained after the end of the project. Throughout the project, only dummy data will be used.
IDIADA	No data will be retained after the end of the project. Throughout the project, only a test dataset will be used.
INTRA	Not applicable.
KUL	Not applicable.
LC	Public dissemination materials, such as newsletters and policy briefs, are archived on the project website and in trusted repositories like Zenodo for at least five years. Personal data are securely deleted in compliance with GDPR. Public dissemination materials are preserved for reference and reuse. Personal data are deleted after project completion. Backups are kept for five years and securely destroyed using GDPR-compliant procedures.
MET	The 2 years dataset will be kept in original format (CSV) until the end of the project.
NORB	The data will be deleted after the end of the project.
QBE	**T4.2** Data are not stored.
	T4.3 Stored on the cloud for the duration of the project, in JSON format and can be deleted through the API or through a request to our DPO or through our website.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS (present status to be updated further during the project)
UMU	Examples used to demonstrate the technology will remain public (e.g., through the project's website/deliverables). No other data needs to be kept.
UoG	No data will be retained after the end of the project. Throughout the project, only synthetic data will be used.
UoM	Upon the completion of the project, the GitHub repositories will remain active. The synthetic data will be disposed in a short period after the completion of the project (within 6 months).
UPRC	All data will remain to the storage area for two years. The format will be same as it will be in the end of the project.

4.1. Please explain what will happen to the data after the end of the project. In doing so, please set out, among others, how/where the data will be kept (e.g., archive, trusted repository for long-term preservation and curation), for how long the data will be kept, in what format they will be kept, how they will be deleted (e.g., according to which technical standard), etc. If not all data will be kept, please explain which data will be kept and how the selection of data for preservation will take place. Please also explain why some data cannot be preserved (e.g., legal or contractual restrictions, storage/budget issues, institutional policies, etc.). Please also specify in your answers for how long any backups will be held and how they will ultimately be destroyed.		
UTH	Data will be deleted after the end of the pilot evaluation.	
VISAR	For the pilot activities specifically, the data used will be dummy data. The handling of this dummy data after the project's end isn't detailed, but it's less critical due to its non-personal nature.	
VTT	**T3.2**	Indefinitely; no personal data included
	T4.1	The data will not be stored after the end of the project. The data will be deleted together with the TANGO pilot infrastructure.
XLAB	Not applicable.	

Annex VI – Making TANGO Data FAIR

Findable

5.1. Will data be identified by a **persistent identifier**? If not, please explain why.		
ABI LAB & NTT	No, the data will be accessible to the banks that will organise them according to internal policies.	
ACV	No, the collected data is provided directly by the data subject. Therefore, there is no persistent identifier linked to the data subject.	
AHOP	Yes.	
ANYS	Yes.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable. No dataset available at the end of the process, but rather ML models and metrics.
	AutoML	Same as above.
	FL	Same as above.
CEA	Yes.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Yes, the encrypted data will be named with a unique transaction number.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR with persistent identifiers.	
FHG	No, the collected data is provided directly IDIADA and are only used for preparing the PoC. Therefore, there is no persistent identifier linked to the raw data.	
FMAKE	Every printed object has a unique persistent name.	
FN	Not applicable.	
FSDE	No, the collected data is provided directly IDIADA and are only used for preparing the PoC. Equivalent data may be reproduced. Therefore, there is no persistent identifier linked to the raw data.	
FUJ_LU	No, the collected data is provided directly by the data subject. Therefore, there is no persistent identifier linked to the data subject.	
IDIADA	No, the collected data is provided directly by IDIADA and are only used for preparing the PoC. Therefore, there is no persistent identifier linked to the raw data.	

5.1. Will data be identified by a **persistent identifier**? If not, please explain why.		
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Public dissemination materials are assigned persistent identifiers, such as DOIs, when deposited in repositories like Zenodo.	
MET	Not yet defined.	
NORB	No, this is handled by the nature of DnAMS and nature of data ownership methodology of our tool.	
QBE	Yes.	
SQD	Not applicable—SQD will not use its own proprietary datasets. During the TANGO project, SQD will assist the RIAS demonstrator with their Dataset Management, including but limited to preparing specific RIAS-TANGO Project datasets. This regards EU Data legislation Compliance, meaning separating specific TANGO Datasets, adding persistent identifiers, implementing standard metadata access protocols, and, if necessary, transforming data for compliance with FAIR principles (present status to be updated further during the project).	
UMU	No, the dataset will not be shared as a whole, it will just serve as examples along documentation of the technical solutions.	
UoG	No, the collected data are provided directly by the data subject. Therefore, there is no persistent identifier linked to that data subject.	
UoM	UoM data are available in GitHub. No personal data shared.	
UPRC	No. Data are free form text collected during interviews and evaluation process.	
UTH	No. Since data will be destroyed after the end of the project, there is therefore no need.	
VISAR	Only dummy data is used.	
VTT	**T3.2**	No. The data has been collected already in 2011, only used in TANGO.
	T4.1	No. Not meant to be accessed by others.
XLAB	Not applicable.	

5.2. Will rich metadata be provided to allow discovery? What metadata will be created? Will disciplinary or general standards be followed, and, if so, which? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.		
ABI LAB & NTT	No, the data will be accessible and managed entirely by the banks.	
ACV	No metadata is created.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.

5.2. Will rich metadata be provided to allow discovery? What metadata will be created? Will disciplinary or general standards be followed, and, if so, which? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.		
	MLOps	No. Metadata will be available according to the MLFlow tool and the experiments searchable within it. Metrics will be defined according to the best practice of data science at will by the user (e.g., precision, recall, etc).
	AutoML	Same as above.
	FL	Not applicable. Metadata could be available to the data scientist in each Pilot, but locally.
CEA	Metadata is composed of current date, implementation targeted, and countermeasure parameters (if any), the number of measurements, number of samples per measurements, the plaintexts used, or any external parameter needed to run the targeted program, picoscope sampling rate, picoscope quantization (number of bits to encode the samples), and the analogue amplification factor of the electromagnetic measurements.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	No.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR and thus take decisions on metadata structure and standards followed.	
FHG	No metadata is created.	
FMAKE	The labels are included in the h5-file. A scientific paper has been published that explains the dataset in more detail: https://doi.org/10.1016/j.addlet.2023.100161. The paper also explains the structure of the data.	
FN	Not applicable.	
FSDE	No metadata is created.	
FUJ_LU	No metadata is created.	
IDIADA	No metadata is created.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Rich metadata are created to allow discovery. Metadata include titles, authors, project information, keywords, descriptions, dates, formats, and version numbers.	
MET	To be discussed with related partners.	
NORB	No, discovery will not be provided.	
QBE	Yes.	
SQD	Yes, all definitions still open at this point (present status to be updated further during the project).	

5.2. Will rich metadata be provided to allow discovery? What metadata will be created? Will disciplinary or general standards be followed, and, if so, which? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.		
UMU	No, although the documents themselves may include metadata (e.g., following W3C's Verifiable Credentials standard, fields like expiration, issuer etc.).	
UoG	No metadata is created.	
UoM	UoM data are available in GitHub. No personal data shared.	
UPRC	Not applicable.	
UTH	Structural metadata will only be used to group data based on device operational parameters.	
VISAR	Not applicable (dummy data).	
VTT	**T3.2**	Not applicable.
	T4.1	No.
XLAB	Not applicable.	

5.3. Will search keywords be provided in the metadata to optimize the possibility for discovery and potential reuse?		
ABI LAB & NTT	No.	
ACV	Not applicable.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable. Only within the MLFlow locally.
	AutoML	Same as above.
	FL	Same as above.
CEA	Countermeasure type and its parameters, and the implementation targeted.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	No.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR and thus having search keywords.	

5.3. Will search keywords be provided in the metadata to optimize the possibility for discovery and potential reuse?		
FHG	Not applicable.	
FMAKE	No, but the data can be found through Google or via the scientific publication.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	No keywords are included in the metadata.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Metadata include relevant search keywords to optimize discoverability and potential reuse.	
MET	No.	
NORB	No.	
QBE	No.	
SQD	Yes, keywords will be supplied by Data Classification.	
UMU	No.	
UoG	No keywords are included in the metadata.	
UoM	No.	
UPRC	Not applicable.	
UTH	No.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	No.
XLAB	Not applicable.	

5.4. What naming conventions do you follow?		
ABI LAB & NTT	The fields are marked with feature_1, feature_2, … , feature_n.	
ACV	Not applicable.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.

5.4. What naming conventions do you follow?		
	MLOps	Not applicable. Only within the MLFlow locally.
	AutoML	Same as above.
	FL	Same as above.
CEA	A folder hierarchy is used, with folders names containing the countermeasures/implementation information.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	The encrypted file will be combined with a random transaction number.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	Not applicable.	
FMAKE	The naming convention is explained on the public website. The naming convention of the file follows the printing layer index.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	No format for the metadata.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	We use clear, descriptive filenames including project acronym, WP number, content description, and version number (e.g., TANGO_WP8_PolicyBrief_v1.0.pdf).	
MET	Not applicable.	
NORB	Not applicable.	
QBE	Based on the data types.	
SQD	Not yet defined.	
UMU	Following common URI, Verifiable Credentials, VPs, DID conventions.	
UoG	No format for the metadata.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	Descriptive file names related to category of log files and relevant timestamp.	
VISAR	Not applicable.	

5.4. What naming conventions do you follow?		
VTT	**T3.2**	Not applicable.
	T4.1	No naming conventions.
XLAB	Not applicable.	

5.5. Do you provide clear version numbers?		
ABI LAB & NTT	No.	
ACV	Not applicable.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	MLFlow keeps track of ML model and experiments versioning. Those are associated to the artifacts in each iteration of the experiments.
	AutoML	Not applicable.
	FL	Not applicable.
CEA	The naming convention will clearly separate different sets of measurements.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	No.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	Not applicable.	
FMAKE	The dataset included a version number.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	Not applicable.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	

5.5. Do you provide clear version numbers?		
LC	Yes. All documents and datasets include clear version numbers in filenames and metadata.	
MET	No.	
NORB	Not applicable.	
QBE	No.	
SQD	Yes, version numbers are a standard dataset classification factor.	
UMU	No.	
UoG	Not applicable.	
UoM	Yes, TornadoVM releases have clear version numbers.	
UPRC	Yes.	
UTH	There is no need.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	No version numbers.
XLAB	Not applicable.	

5.6. Will be metadata be offered in a way that can be harvested and indexed?		
ABI LAB & NTT	No.	
ACV	No.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable. Only locally, not shared.
	AutoML	Same as above.
	FL	Same as above.
CEA	Yes, in a txt file.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	No.	
ECO	Not applicable.	

5.6. Will be metadata be offered in a way that can be harvested and indexed?		
EGI	Not applicable.	
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR and thus providing details on harvesting and indexing.	
FHG	Not applicable.	
FMAKE	At the moment, the metadata is included in the dataset and is not indexed.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	No.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Metadata are deposited in trusted repositories such as Zenodo.	
MET	To be discussed.	
NORB	Not applicable.	
QBE	It will be in human readable format.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be curated previously to being used in the TANGO Project, with the goals of supplying anonymized, homogenous, datasets to the required TANGO activities that require harvesting and indexing (present status to be updated further during the project).	
UMU	No.	
UoG	No.	
UoM	No.	
UPRC	Not applicable.	
UTH	Not applicable.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	No.
XLAB	Not applicable.	

Accessible

Data

5.7. Will all data be made openly available?		
ABI LAB & NTT	No.	
ACV	No.	
AHOP	No, datasets provided in TANGO project are only available for Pilot purposes.	
ANYS	No, datasets provided in TANGO project are only available for Pilot purposes.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable. Only examples will be made openly available, but not Pilots' data, e.g. in technical documentation of the solutions.
	AutoML	Same as above.
	FL	Same as above.
CEA	The first dataset is openly available.	
CESGA	Not applicable.	
DBC	No.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR.	
FHG	No.	
FMAKE	Yes.	
FN	Not applicable.	
FSDE	No.	
FUJ_LU	No.	
IDIADA	No.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Not all data are made openly available. Public dissemination materials are openly accessible, while personal data (e.g., participant lists, newsletter subscriptions) are not shared due to GDPR requirements.	

MET	No.
NORB	No.
QBE	No.
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).
UMU	Only as examples—e.g., in technical documentation of the solutions.
UoG	No.
UoM	Yes.
UPRC	Yes, to project partners. Processed information will be publicly available in D2.2, D7.1, D7.2.
UTH	Not applicable.
VISAR	No.
VTT	**T3.2** Not applicable.
	T4.1 No.
XLAB	Not applicable.

5.8. If certain datasets cannot be shared (or need to be shared under restrictions), please explain (i) what restrictions will apply (including identifying the data that will not be made available), and (ii) why these restrictions apply. In doing so, please clearly separate legal and contractual reasons from voluntary restrictions.		
ABI LAB & NTT	All the data of the banks cannot be shared.	
ACV	No dataset will be shared outside the organization that collects the data from the data subject.	
AHOP	No input provided.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable to the components. This is under discussion amongst Pilots' partners, but most probably is not going to be shared.
	AutoML	Same as above.
	FL	Same as above.
CEA	The second dataset is currently closed. Its use is deemed of low interest for the community; however, upon request we may open it.	
CESGA	Not applicable.	

5.8. If certain datasets cannot be shared (or need to be shared under restrictions), please explain (i) what restrictions will apply (including identifying the data that will not be made available), and (ii) why these restrictions apply. In doing so, please clearly separate legal and contractual reasons from voluntary restrictions.	
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR.
FHG	No dataset will be shared outside the organization that collects the data from the data subject.
FMAKE	Not applicable, as datasets have been shared.
FN	Not applicable.
FSDE	No dataset will be shared outside the organization that collects the data from the data subject.
FUJ_LU	No dataset will be shared outside the organization that collects the data from the data subject.
IDIADA	No dataset will be shared outside the organization that collects the data from the data subject.
INTRA	Not applicable.
KUL	Not applicable.
LC	Personal data cannot be shared due to legal (GDPR) restrictions protecting privacy. Only anonymised, aggregated data are shared in dissemination reports. Internal working documents are restricted to consortium members due to confidentiality agreements.
MET	Not applicable.
NORB	Data is owned by the users of the pilot and not the TANGO members.
QBE	These are biometric and behavioural data.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS are still under definition but personal data collection, transformation, and sharing will respect Portuguese government issued legislation regarding GPDR principles. SQD must inform that collection of data such as racial/ethnical/political/religion/association with unions, and others belonging to individuals are specifically forbidden under Portuguese Law, therefore those information classes are not inquired or collected, and will not be part of any TANGO dataset. Please refer to https://dre.pt/dre/en/detail/act/58-2019-123815982.
UMU	Not applicable.
UoG	No dataset will be shared outside the organization that collects the data from the data subject.
UoM	Not applicable.

5.8. If certain datasets cannot be shared (or need to be shared under restrictions), please explain (i) what restrictions will apply (including identifying the data that will not be made available), and (ii) why these restrictions apply. In doing so, please clearly separate legal and contractual reasons from voluntary restrictions.		
UPRC	As long as the project is running or stated by the CO.	
UTH	No, data will be only available to UTH and RIAS.	
VISAR	Not applicable.	
VTT	**T3.2**	The project which collected the Legacy dataset ended in 2013 and did not publish the dataset.
	T4.1	Dataset cannot be shared as it contains personal identity documents not collected for research use.
XLAB	Not applicable.	

5.9. How will data be made accessible (e.g., by deposition in repository)?	
ABI LAB & NTT	Not applicable.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Deposit on Zenodo, along with publication.
CESGA	Not applicable.
DBC	Not applicable.
DUT	The data is indexed by unique transaction number and decrypted with keys. The data is stored in local server and accessed with API.
ECO	Not applicable.
EGI	Not applicable.
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR being accessible in public repositories.
FHG	Not applicable.
FMAKE	Through a custom website and data hosted by Juelich; accessible through https.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	Not applicable.
IDIADA	Not applicable.
INTRA	Not applicable.

5.9. How will data be made accessible (e.g., by deposition in repository)?		
KUL	Not applicable.	
LC	Public dissemination materials are made accessible via deposition in trusted repositories (e.g., Zenodo) and the TANGO project website.	
MET	Not applicable.	
NORB	Data will not be made accessible.	
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).	
UMU	Accompanying documentation—e.g., in code repository or the project's deliverables.	
UoG	Not applicable.	
UoM	GitHub access, and docker images available in DockerHub.	
UPRC	At the time of a publication.	
UTH	Not publicly available.	
VISAR	Data will be stored inside the TANGO component Fides, specifically in its distributed legally compliant data backend DRILL, and that the Fides API document (WIP) will provide information on available APIs/types of queries for data retrieval.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.10. Please explain when will data be made available to third parties (non-TANGO partners) (e.g., as soon as possible after the data collection, at the end of the project, at the time of a publication, etc.?). If an embargo is applied, please explain why and for how long it will apply, keeping in mind that research data should be made available as soon as possible.		
ABI LAB & NTT	The data is owned by the banks and will remain in the banks' on-premises environments.	
ACV	Not applicable.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable.

5.10. Please explain when will data be made available to third parties (non-TANGO partners) (e.g., as soon as possible after the data collection, at the end of the project, at the time of a publication, etc.?). If an embargo is applied, please explain why and for how long it will apply, keeping in mind that research data should be made available as soon as possible.		
	AutoML	Same as above.
	FL	Same as above.
CEA	First dataset was released at the time of the associated publication.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	Not applicable.	
FMAKE	The data is already available to non-TANGO partners, since it was an outcome of a past project which is already finished.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	Not applicable.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Public dissemination materials are made available as soon as they are finalized and approved, typically shortly after publication or project events. No embargo is applied.	
MET	Not applicable.	
NORB	No data will be made available to the third parties.	
QBE	The data collected are special category of personal data. Consent has been received for training the models but there is no consent of sharing the data outside of QBE organization.	
SQD	Not applicable—SQD will not use its own proprietary datasets, RIAS will have full control of their own Datasets, transfer of RIAS Data to TANGO partners will be done as required by the TANGO SOW activities.	
UMU	At the time of publication of documents and technical descriptions.	
UoG	Not applicable.	
UoM	Not applicable.	

5.10. Please explain when will data be made available to third parties (non-TANGO partners) (e.g., as soon as possible after the data collection, at the end of the project, at the time of a publication, etc.?). If an embargo is applied, please explain why and for how long it will apply, keeping in mind that research data should be made available as soon as possible.		
UPRC	Data are stored in document files (MS Word and pdf).	
UTH	Not publicly available.	
VISAR	At the end of the project.	
VTT	**T3.2**	If requested, publishing may be considered.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.11. What methods or software tools are needed to access the data?		
ABI LAB & NTT	Not applicable.	
ACV	Text editor.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Provided by the components locally.
	AutoML	Same as above.
	FL	Same as above.
CEA	NumPy library, python3, and text editors.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	User's private API, SEDSS component, and CP-ABE.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	Not applicable.	
FMAKE	The data can be downloaded using https; so there are many browsers/tools available. Any package that supports hdf5 will be able to read the data. There are free and open-source Python packages available. Example code is available here: https://github.com/Flanders-Make-vzw/RAISE_LPBF_Laser_benchmark.	
FN	Not applicable.	

5.11. What methods or software tools are needed to access the data?	
FSDE	Not applicable.
FUJ_LU	PAT provides a secure protocol for the organization collecting the data to access it.
IDIADA	Not applicable.
INTRA	Not applicable.
KUL	Not applicable.
LC	No special software is required. All data are in standard formats (PDF, MP4, PNG, CSV) accessible by common tools.
MET	Not yet defined.
NORB	Not applicable.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).
UMU	Depending on the medium, browser, or text readers.
UoG	Django secure authentication process.
UoM	GitHub.
UPRC	Not applicable
UTH	https requests.
VISAR	Data will be stored inside the TANGO component Fides, specifically in its distributed legally compliant data backend DRILL, using internal Fides format.
VTT	**T3.2** Not applicable.
	T4.1 Data will not be made accessible.
XLAB	Not applicable.

5.12. Will the data be accessible through a free and standardised access protocol?		
ABI LAB & NTT	No.	
ACV	No.	
AHOP	Not applicable.	
ANYS	Not applicable.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable.

5.12. Will the data be accessible through a free and standardised access protocol?		
	AutoML	Same as above.
	FL	Same as above.
CEA	The data on Zenodo is downloadable easily from everyone.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	No.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets will be provided as open and FAIR in a standardized access protocol.	
FHG	No, the data will not be accessible through a free and standardized access protocol. Access to the data will be strictly controlled and limited to authorized personnel within the organization that collects the data, ensuring compliance with data protection regulations and maintaining data security.	
FMAKE	See above, https is used.	
FN	Not applicable.	
FSDE	No, the data will not be accessible through a free and standardized access protocol. Access to the data will be strictly controlled and limited to authorized personnel within the organization that collects the data, ensuring compliance with data protection regulations and maintaining data security.	
FUJ_LU	No, the data will not be accessible through a free and standardized access protocol. Access to the data will be strictly controlled and limited to authorized personnel within the organization that collects the data, ensuring compliance with data protection regulations and maintaining data security.	
IDIADA	No, the data will not be accessible through a free and standardized access protocol. Access to the data will be strictly controlled and limited to authorized personnel within the organization that collects the data, ensuring compliance with data protection regulations and maintaining data security.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Data are accessible via free, standardized protocols such as HTTPS and repository APIs.	
MET	Not yet defined.	
NORB	Not applicable.	
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms,	

5.12. Will the data be accessible through a free and standardised access protocol?		
	processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).	
UMU	Yes.	
UoG	Access to the data will be strictly controlled and limited to authorised personnel within the organisation that collects the data.	
UoM	The licenses are permissive. The TornadoVM API is licensed under Apache 2.0.	
UPRC	Not applicable.	
UTH	Yes, through https.	
VISAR	Access to the data, particularly within the Fides system, will be mediated through APIs.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.13. If there are restrictions on use of the data, how will access be provided to the data both during and after the end of the project?		
ABI LAB & NTT	The data will be accessed directly on the banks' environments.	
ACV	No access to the data after the project.	
AHOP	Not applicable.	
ANYS	No input provided.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Not applicable. Depends on the Pilots.
	AutoML	Same as above.
	FL	Same as above.
CEA	No restrictions.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	No access to the data after the project.	

5.13. If there are restrictions on use of the data, how will access be provided to the data both during and after the end of the project?		
FMAKE	Without any other agreement, license CC BY-NC-4.0 applies.	
FN	Not applicable.	
FSDE	No access to the data after the project.	
FUJ_LU	No access to the data after the project.	
IDIADA	No access to the data after the project.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Restricted datasets (personal data) are accessible only to authorized LIC personnel and are not publicly shared. Aggregated and anonymised results are shared internally during the project and archived publicly after the project ends.	
MET	Not yet defined.	
NORB	Data will be available to authorized users during the Pilot only.	
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).	
UMU	Not applicable.	
UoG	No access to the data after the project.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	Authorization mechanism embedded to database system (Influx DB).	
VISAR	No restrictions for the dummy data.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.14. How will the identity of the person accessing the data be ascertained?	
ABI LAB & NTT	In charge of banks.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	Not applicable.

5.14. How will the identity of the person accessing the data be ascertained?	
ATOS	Not applicable.
CEA	Access control on CEA servers, or for the data sent through file senders, ciphering it.
CESGA	Not applicable.
DBC	Not applicable.
DUT	By authentication mechanism of TANGO architecture.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	The data shall be accessed only on specific request and with agreement by IDIADA.
FMAKE	Anonymous access is enabled; as such, no identification is needed.
FN	Not applicable.
FSDE	The data shall be accessed only on specific request and with agreement by IDIADA.
FUJ_LU	Not applicable.
IDIADA	The data shall be accessed only on specific request and with agreement by IDIADA.
INTRA	Not applicable.
KUL	Not applicable.
LC	For restricted internal data, access is controlled through role-based permissions within Lisbon Council's secure systems or the ones of the project. Public data are openly accessible without authentication.
MET	Not yet defined.
NORB	Using ID check mechanism such as eID if available in the country of a pilot or 3rd party ID verification providers like Onfido, Jumio etc.
QBE	Not applicable.
SQD	N/A—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project)
UMU	Not applicable.
UoG	Not applicable.
UoM	Not applicable.
UPRC	User account for using the project's OwnCloud.
UTH	Through authentication mechanism.

5.14. How will the identity of the person accessing the data be ascertained?		
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Not applicable.
XLAB	Not applicable.	

5.15. Is there a need for a data access committee (e.g., to evaluate/approve access requests to personal/sensitive data)?	
ABI LAB & NTT	In charge of banks.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	No.
CESGA	No.
DBC	Not applicable.
DUT	No.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	No.
FMAKE	No.
FN	Not applicable.
FSDE	No need.
FUJ_LU	No.
IDIADA	No.
INTRA	Not applicable.
KUL	Not applicable.
LC	No. There is no need for a data access committee as personal data are not shared externally.
MET	To be discussed in case of data sharing.
NORB	No, the data is not available to 3rd parties.
QBE	Not applicable.

5.15. Is there a need for a data access committee (e.g., to evaluate/approve access requests to personal/sensitive data)?		
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).	
UMU	No.	
UoG	No.	
UoM	No.	
UPRC	No.	
UTH	No.	
VISAR	No.	
VTT	**T3.2**	Not applicable.
	T4.1	Not applicable.
XLAB	Not applicable.	

Metadata

5.16. Will metadata be made openly available and licenced under a public domain dedication CCO? If not, please explain why.	
ABI LAB & NTT	No, the data is owned by the banks and will remain in the banks' on-premises environments
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	No.
CEA	The metadata will be described in the publications that present the obtained results.
CESGA	Not applicable.
DBC	Not applicable.
DUT	No.
ECO	Not applicable.
EGI	Not applicable.
EXUS	EXUS as a technical partner processes the data provided by end users and cannot take any decision on whether those datasets or metadata information will be provided as open and FAIR.
FHG	No, as no metadata are collected.
FMAKE	Metadata and data are publicly available.

5.16. Will metadata be made openly available and licenced under a public domain dedication CCO? If not, please explain why.		
FN	Not applicable.	
FSDE	No, as no metadata are collected.	
FUJ_LU	No.	
IDIADA	No, as no metadata are collected.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Metadata are openly available and licensed when needed.	
MET	Not applicable.	
NORB	No.	
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project)	
UMU	Not applicable.	
UoG	No.	
UoM	No metadata created.	
UPRC	Not applicable.	
UTH	No. Data and metadata is only intended for use during the evaluation of the pilots.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.17. Will metadata contain information to enable the user to access the data?	
ABI LAB & NTT	No.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Only in case that the data is published with the publication.
CESGA	Not applicable.

5.17. Will metadata contain information to enable the user to access the data?	
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Not applicable.
FMAKE	Yes.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	No.
IDIADA	Not applicable.
INTRA	Not applicable.
KUL	Not applicable.
LC	Yes. Metadata include links to access the data and details such as format, description, licensing, and persistent identifiers.
MET	Not applicable.
NORB	Not applicable.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).
UMU	Not applicable.
UoG	No.
UoM	Not applicable.
UPRC	Not applicable.
UTH	No.
VISAR	No.
VTT	**T3.2** Not applicable.
	T4.1 Data will not be made accessible.
XLAB	Not applicable.

5.18. For how long will the data remain available and findable? Will metadata be guaranteed to remain available after the data is no longer available? For how long metadata will remain available?	
ABI LAB & NTT	Not applicable.
ACV	No.
AHOP	No input provided.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Whenever Zenodo remains available for opened data. All metadata published will survive independently, as part of the publication.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	The test dataset will be stored only for the duration of the TANGO project.
FMAKE	The data will be hosted for several years. The metadata is described in the scientific open publication and will remain available.
FN	Not applicable.
FSDE	The test dataset will be stored only for the duration of the TANGO project.
FUJ_LU	Not applicable.
IDIADA	The test dataset will be stored only for the duration of the TANGO project.
INTRA	Not applicable.
KUL	Not applicable.
LC	Public dissemination materials and their metadata remain available for at least five years after the end of the project. Metadata persist even if data are removed, ensuring long-term findability.
MET	Both will remain available.
NORB	Data will be deleted right after the Pilot.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project)
UMU	Not applicable.
UoG	Not applicable.

5.18. For how long will the data remain available and findable? Will metadata be guaranteed to remain available after the data is no longer available? For how long metadata will remain available?		
UoM	Not applicable.	
UPRC	As long as the project is running.	
UTH	During pilot evaluation.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.19. Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g., in open-source code)?	
ABI LAB & NTT	No, there will be no possibility to access or read the data.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Widely accepted formats like txt or NumPy can be read by external programs easily, the procedures are already largely documented on the web.
CESGA	Not applicable.
DBC	Not applicable.
DUT	No.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	No.
FMAKE	A git repository with example code is available.
FN	Not applicable.
FSDE	No need.
FUJ_LU	No.
IDIADA	No need.
INTRA	Not applicable.
KUL	Not applicable.

5.19. Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g., in open-source code)?	
LC	No special software is needed. All materials are provided in standard, widely accessible formats. No proprietary software is required.
MET	Not yet defined.
NORB	No.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project)
UMU	Not applicable.
UoG	No.
UoM	Documentation is available in GitHub as readme: https://github.com/TANGO-EU-PROJECT/docker-tornadovm/blob/test/cicd/RUN_DYNAMIC_INTELLIGENT_TASK.md Further blogs may emerge either in the TANGO website, or the TornadoVM website.
UPRC	Not applicable.
UTH	Yes, UTH provide the appropriate guidelines.
VISAR	Not applicable.
VTT	**T3.2** Not applicable.
	T4.1 Data will not be made accessible.
XLAB	Not applicable.

Repository

5.20. Where will the data and associated metadata, documentation and code be deposited? Will they be deposited in a trusted repository? If so, please indicate the repository chosen.	
ABI LAB & NTT	No, the data is owned by the banks and will remain in the banks' on-premises environments.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	On Zenodo and CEA servers.
CESGA	Not applicable.
DBC	Not applicable.
DUT	No metadata generated. The data is stored in user's local server, which is trusted from users' side.

5.20. Where will the data and associated metadata, documentation and code be deposited? Will they be deposited in a trusted repository? If so, please indicate the repository chosen.	
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Documentation on the use case and evaluation will be stored in the TANGO repository. The Docker image of the TSM will be deployed in the TANGO integration environment.
FMAKE	No, the data size is too large and is showed on a particular website.
FN	Not applicable.
FSDE	Documentation on the use case and evaluation will be stored in the TANGO repository. The Docker image of the TSM will be deployed in the TANGO integration environment.
FUJ_LU	Not applicable.
IDIADA	Documentation on the use case and evaluation will be stored in the TANGO repository. The Docker image of the TSM will be deployed in the TANGO integration environment.
INTRA	Not applicable.
KUL	Not applicable.
LC	Data and metadata are deposited in trusted repositories such as Zenodo and made available via the TANGO website.
MET	No.
NORB	Data will be stored in Fides DRILL. The code will not be deposited because of the closed nature of the code.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).
UMU	Alongside the documentation—e.g., in the project's website.
UoG	No.
UoM	GitHub repositories, identified as dataset above.
UPRC	Atos OwnCloud and maybe Zenodo.
UTH	GitHub.
VISAR	Data will be stored inside the TANGO component Fides, specifically in its distributed legally compliant data backend DRILL.
VTT	**T3.2** Not applicable.
	T4.1 Data will not be made accessible.
XLAB	Not applicable.

5.21. Have you explored appropriate arrangements with the identified repository where your data will be deposited?	
ABI LAB & NTT	No.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	No.
CESGA	Not applicable.
DBC	Not applicable.
DUT	No.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	The project CA shall cover this.
FMAKE	No.
FN	Not applicable.
FSDE	The project CA shall cover this.
FUJ_LU	No.
IDIADA	The project CA shall cover this.
INTRA	Not applicable.
KUL	Not applicable.
LC	Yes. LC follows the standard policies of Zenodo for data deposition.
MET	No.
NORB	No, the data will not be deposited.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project).
UMU	Not applicable.
UoG	No.
UoM	Yes.
UPRC	Yes.

5.21. Have you explored appropriate arrangements with the identified repository where your data will be deposited?		
UTH	Yes.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.22. Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?	
ABI LAB & NTT	No.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	It depends on the repository.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Subject to user's local environment. SEDSS do not require it.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Not applicable.
FMAKE	Not applicable.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	No.
IDIADA	Not applicable.
INTRA	Not applicable.
KUL	Not applicable.
LC	Yes. Zenodo assigns a DOI to each dataset and ensures persistent resolution.
MET	No.
NORB	Not applicable.

5.22. Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?		
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be kept and stored by RIAS, and access mechanisms, processual and others will be implemented according with RIAS requirements for regulatory compliance (present status to be updated further during the project)	
UMU	Not applicable.	
UoG	No.	
UoM	No.	
UPRC	Yes.	
UTH	Yes.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

Interoperable

5.23. Are the produced data interoperable, allowing data exchange and reuse between researchers, institutions, organisations, etc. (in other words, do they adhere to standards for formats, in line, as much as possible, with available (open) software applications, and do they enable re-combinations with different datasets from different origins)?		
ABI LAB & NTT	No.	
ACV	No input provided.	
AHOP	Not applicable.	
ANYS	All datasets are interoperable thanks to FIWARE Context Broker.	
ATOS	**ID**	Yes, it aligns with W3C SSI standards and the integration profile specified by EBSI Conformance Guidelines (VCs,VPs, DIDs) and JSON-LD.
	SP	No, the ABE keys are implementation-specific binary data. Likewise, the encrypted data is encrypted with implementation specific encryption and is similarly not interoperable with other ABE systems.
	MLOps	Not applicable. Standard ML models.
	AutoML	Same as above.
	FL	Same as above.
CEA	Yes, thanks to the data storage format used.	

5.23. Are the produced data interoperable, allowing data exchange and reuse between researchers, institutions, organisations, etc. (in other words, do they adhere to standards for formats, in line, as much as possible, with available (open) software applications, and do they enable re-combinations with different datasets from different origins)?		
CESGA	All datasets are interoperable thanks to FIWARE Context Broker.	
DBC	Not applicable.	
DUT	No. Data are all encrypted.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	Not applicable.	
FMAKE	Yes, standard hdf5 format.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	PAT uses JSON formatted data, which is a well-known standard way of structuring data.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Data are produced in interoperable formats (PDF, CSV, MP4) compatible with open software applications and suitable for reuse.	
MET	No.	
NORB	No, the data will not be interoperable.	
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets. During the TANGO project SQD will assist the RIAS demonstrator with their Dataset Management, including but limited to preparing specific RIAS-TANGO Project datasets, in what regards EU Data legislation Compliance, this meaning separating specific TANGO Datasets, adding persistent identifiers, implementing standard metadata access protocols, and if necessary transforming data for compliance with FAIR principles (present status to be updated further during the project).	
UMU	**ID**	Yes, it aligns with W3C SSI standards (VCs,VPs, DIDs) and JSON-LD.
	PEP	Yes, it aligns with JSON format.
	SP	Same as above.
UoG	No.	
UoM	Yes.	

5.23. Are the produced data interoperable, allowing data exchange and reuse between researchers, institutions, organisations, etc. (in other words, do they adhere to standards for formats, in line, as much as possible, with available (open) software applications, and do they enable re-combinations with different datasets from different origins)?		
UPRC	Yes.	
UTH	No.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.24. What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and reuse within and across disciplines? Will you follow community-endorsed interoperability best practices? If so, which ones?		
ABI LAB & NTT	Not applicable.	
ACV	No.	
AHOP	Not applicable.	
ANYS	NGSI-V2.	
ATOS	**ID**	W3C's Verifiable Credentials, which is the leading document for SSI solutions.
	SP	Use of JSON for sticky policies, including common fields in the literature like id, schemas.
	MLOps	Input data according to the needs of the tools. Usage according to standard ML practices.
	AutoML	Same as above.
	FL	Same as above.
CEA	We will use standard data storage format and will provide the necessary information to process the data and understand what it is, associating a publication with it alongside its metadata. The vocabulary will be standard to the side-channel community.	
CESGA	No input provided.	
DBC	Not applicable.	
DUT	The data identifier is encoded by UTF-8. The data is encrypted, which is not readable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	Not applicable.	

5.24. What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and reuse within and across disciplines? Will you follow community-endorsed interoperability best practices? If so, which ones?		
FMAKE	hdf5 format.	
FN	Not applicable.	
FSDE	Not applicable.	
FUJ_LU	No.	
IDIADA	Not applicable.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Data formats are open and widely used (PDF, CSV). Best practices for interoperability are applied to allow exchange and reuse.	
MET	To be discussed.	
NORB	Not applicable.	
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets. During the TANGO project, SQD will assist the RIAS demonstrator with their Dataset Management, including but limited to preparing specific RIAS-TANGO Project datasets, in what regards EU Data legislation Compliance. This means separating specific TANGO Datasets, adding persistent identifiers, implementing standard metadata access protocols, and if necessary transforming data for compliance with FAIR principles (present status to be updated further during the project).	
UMU	**ID**	W3C's Verifiable Credentials, which is the leading document for SSI solutions
	PEP	Use of JSON including common fields in the literature like id, schemas etc.
	SP	Same as above.
UoG	No.	
UoM	Community-driven interoperability. Please create a drop-down in order to identify the practices.	
UPRC	Not applicable.	
UTH	None, there is no need as data in intended only for use in Pilot evaluation.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.25. If it is unavoidable that you use uncommon or generate project-specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow re-using, refining or extending them?	
ABI LAB & NTT	Not applicable.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Yes.
CESGA	No input provided.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Not applicable.
FMAKE	The data structure is embedded in the hdf5 format and also documented on the public website.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	No.
IDIADA	Not applicable.
INTRA	Not applicable.
KUL	Not applicable.
LC	Not applicable. No project-specific ontologies or uncommon vocabularies are used.
MET	To be discussed.
NORB	Not applicable.
QBE	Not applicable.
SQD	Yes, specific Ontology mappings will be provided as per relevant user requirements.
UMU	If needed, mappings will be published along with the research on the technologies that demanded it.
UoG	No.
UoM	Not applicable.
UPRC	Not applicable.

5.25. If it is unavoidable that you use uncommon or generate project-specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow re-using, refining or extending them?		
UTH	No, there is no need as data is intended only for use in Pilot evaluation.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.26. Will your data include qualified references to other data (e.g., data from your project or datasets from previous research)?	
ABI LAB & NTT	No.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	No.
CESGA	No input provided.
DBC	Not applicable.
DUT	No.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	No other data within TANGO used. The API of the TSM is documented within the project's deliverables.
FMAKE	No reference to other data.
FN	Not applicable.
FSDE	No other data within TANGO used. The API of the TSM is documented within the project's deliverables.
FUJ_LU	No.
IDIADA	No other data within TANGO used. The API of the TSM is documented within the project's deliverables.
INTRA	Not applicable.
KUL	Not applicable.
LC	Yes. Dissemination materials reference related TANGO datasets and prior research where appropriate.

5.26. Will your data include qualified references to other data (e.g., data from your project or datasets from previous research)?		
MET	No.	
NORB	Not applicable.	
QBE	Not applicable.	
SQD	N/A—SQD will not use its own proprietary datasets, Datasets to be handled from RIAS will be curated previously to being used in the TANGO Project, with the goals of supplying anonymized, homogenous, hindsight datasets to the required TANGO activities (present status to be updated further during the project).	
UMU	No.	
UoG	No.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	No.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

Reusable

5.27. How will you provide documentation needed to validate data analysis and facilitate data reuse (e.g., "readme" files with information on methodology, codebooks, data cleaning, analysis, variable definitions, units of measurement, etc.)?		
ABI LAB & NTT	Not applicable.	
ACV	No.	
AHOP	No.	
ANYS	No.	
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	Technical documentation in deliverables. Docker containers including readme files for the software components.
	AutoML	Same as above.
	FL	Same as above.

5.27. How will you provide documentation needed to validate data analysis and facilitate data reuse (e.g., "readme" files with information on methodology, codebooks, data cleaning, analysis, variable definitions, units of measurement, etc.)?	
CEA	README files are available alongside the data and describe how to read it. The results of our analysis on the data are made available as well though the publications, which allows to validate further data analysis by comparing the outcomes.
CESGA	No input provided.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	The privacy preserving module as well as the use case shall be documented within the TANGO deliverables.
FMAKE	Documentation is available on the website and in a public scientific publication.
FN	Not applicable.
FSDE	The TSM module as well as the use case shall be documented within the TANGO deliverables.
FUJ_LU	Not applicable.
IDIADA	The TSM module as well as the use case shall be documented within the TANGO deliverables.
INTRA	Not applicable.
KUL	Not applicable.
LC	Documentation such as branding guidelines, communication plans, and a "readme" file describing the structure, purpose, and licensing of each dataset is included.
MET	Not yet defined.
NORB	Not applicable.
QBE	Not applicable.
SQD	Not applicable—SQD will not use its own proprietary datasets, Datasets from RIAS still under definition (present status to be updated further during the project).
UMU	Shared documents will include descriptions of the data, and in technical documentations the formats, designs and application of the documents will be described (methodology, fields used, and purpose etc.).
UoG	Documentation will be provided to ensure transparency, validation of data analysis, and support for data reuse. This will include "readme" files outlining the methodology, data structures, and analytical processes.
UoM	Yes, readme files.
UPRC	No input provided.
UTH	We will provide instructions, guidelines, and readme files to offer the sufficient amount of information that will allow reproducing the same data.

5.27. How will you provide documentation needed to validate data analysis and facilitate data reuse (e.g., "readme" files with information on methodology, codebooks, data cleaning, analysis, variable definitions, units of measurement, etc.)?		
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.28. Will your data be made freely available to the public domain to permit the widest reuse possible? Will your data be licensed using standard reuse licenses? Please explain.	
ABI LAB & NTT	No.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Data uses standard reuse licenses.
CESGA	No, data from Smart Hospitality pilot is only for pilot purposes. Only data from the Experiences provider will be reused.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	No.
FMAKE	It is published under the CC BY-NC 4.0 license.
FN	Not applicable.
FSDE	No.
FUJ_LU	No.
IDIADA	No.
INTRA	Not applicable.
KUL	Not applicable.
LC	Yes. Public dissemination materials are made freely available.
MET	Not yet defined.
NORB	No, the data will not be made freely available for the public domain.

5.28. Will your data be made freely available to the public domain to permit the widest reuse possible? Will your data be licensed using standard reuse licenses? Please explain.	
QBE	No.
SQD	N/A—SQD will not use its own proprietary datasets. Licensing of use for RIAS datasets still under definition (present status to be updated further during the project).
UMU	Examples will be published along with the technical descriptions of the solution, through public deliverables, repositories, or research publications.
UoG	Yes, it will be provided for reuse.
UoM	Yes, the license of the TornadoVM API is permissive (Apache 2.0).
UPRC	Not applicable.
UTH	No, there is no need as data in intended only for use in pilot evaluation.
VISAR	Not applicable.
VTT	**T3.2** Not applicable.
	T4.1 Data will not be made accessible.
XLAB	Not applicable.

5.29. Will the data produced in the project be usable by third parties, in particular after the end of the project? Please provide an overview.	
ABI LAB & NTT	No.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Yes.
CESGA	No.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	No.
FMAKE	Not clear at the moment. It is not expected that partners used the data after the project.
FN	Not applicable.

5.29. Will the data produced in the project be usable by third parties, in particular after the end of the project? Please provide an overview.		
FSDE	No.	
FUJ_LU	No.	
IDIADA	No.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Dissemination materials and associated documentation are usable by third parties, including policymakers, researchers, and stakeholders, after the project ends.	
MET	No.	
NORB	No, the data produced will not be used by 3rd parties and it will be allowed only for the stakeholders and participants of the pilot project.	
QBE	No.	
SQD	Not applicable—SQD will not use its own proprietary datasets, use of RIAS Datasets still under definition (present status to be updated further during the project).	
UMU	**ID**	It will be useful as an exemplification of the technology and its uses.
	PEP	It will be useful to demonstrate the PEP module as used in TANGO, and serve as a reference format for it
	SP	It will be useful to demonstrate the CP-ABE technology as used in TANGO, and serve as a reference format for it.
UoG	Yes, will be made available through GitHub.	
UoM	Yes, the source code of the implementation in the context of T5.3 is available in GitHub (TANGO forks of TornadoVM and docker-tornadovm).	
UPRC	Not applicable.	
UTH	No, there is no need as data is intended only for use in pilot evaluation.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.30. Will the provenance of the data be thoroughly documented using the appropriate standards?	
ABI LAB & NTT	Data management is entirely the responsibility of the owners (the banks).
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.

5.30. Will the provenance of the data be thoroughly documented using the appropriate standards?		
ATOS	**ID**	Not applicable. Identity data is not shared.
	SP	Not applicable. Cryptographic keys stored in ABE Toolset are not shared.
	MLOps	In MLOps, experiments are tracked as well as artifacts and metadata.
	AutoML	Same as above.
	FL	Same as above.
CEA	Yes.	
CESGA	Not applicable.	
DBC	Not applicable.	
DUT	Not applicable.	
ECO	Not applicable.	
EGI	Not applicable.	
EXUS	Not applicable.	
FHG	The use case documentation shall include this.	
FMAKE	No.	
FN	Not applicable.	
FSDE	The use case documentation shall include this.	
FUJ_LU	No.	
IDIADA	The use case documentation shall include this.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Data provenance is documented, including authorship, creation date, and versioning.	
MET	No.	
NORB	Not applicable.	
QBE	Not applicable.	
SQD	Yes.	
UMU	It will be clearly stated that it was generated through the project's work.	
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	No, there is no need as data in intended only for use in Pilot evaluation.	

5.30. Will the provenance of the data be thoroughly documented using the appropriate standards?		
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made accessible.
XLAB	Not applicable.	

5.31. Describe all relevant quality assurance processes.	
ABI LAB & NTT	The data is provided by banks that implement proprietary quality standards.
ACV	No.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Data published was downloaded again to check that it corresponds to the data we analysed to generate our results. This verification is by generating hashes of the 2 datasets.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	D5.1 & D5.2 offer a description for EDAE's related activities.
FHG	The QA processes of TANGO apply.
FMAKE	Not applicable.
FN	Not applicable.
FSDE	The QA processes of TANGO apply.
FUJ_LU	Not applicable.
IDIADA	The QA processes of TANGO apply.
INTRA	Not applicable.
KUL	Not applicable.
LC	LC applies internal quality control procedures, including editorial review, partner feedback, approval workflows, and version control. Data accuracy and consistency are checked before publication.
MET	Not applicable.
NORB	Not applicable.

5.31. Describe all relevant quality assurance processes.		
QBE	Not applicable.	
SQD	Not applicable—SQD will not use its own proprietary datasets. QA processes to be implemented on RIAS Datasets still under definition (present status to be updated further during the project).	
UMU	Not applicable.	
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	UTH will take all the measures to ensure relevance, accuracy, consistency, timeliness, and compliance of the data that will use relying on statistical tools, identifying outliers, checking legal compliance and legislation.	
VISAR	Not applicable.	
VTT	**T3.2**	No, there is no need as data in intended only for use in pilot evaluation.
	T4.1	No, there is no need as data in intended only for use in pilot evaluation.
XLAB	Not applicable.	

Annex VII – Other Research Outputs

6.1. Please identify research outputs other than data that you will generate during the TANGO project. Such outputs may be either digital (e.g., software, workflows, protocols, models, etc.) or physical (e.g., samples, new materials, etc.).	
ABI LAB & NTT	Development of a methodology to approach and mitigate the risk of fraud on banking transactions.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Software for the technical solutions. Atos is the sole owner of Intellectual Property rights for the SSI Agent (solutions are not co-developed with other partners) and Federated Learning framework (based on the FLEVIDEN background and foreground). MLOps and AutoML are open-source bundles tailored to the project's needs. ATOS participates in the definition of the workflows, models and architectural design of the project's results
CEA	Software.
CESGA	Not applicable.
DBC	Not applicable.
DUT	The encryption-decryption workflow will be packed into a standalone plugin, which can be used/deployed to other projects or software in the future.
ECO	Not applicable.
EGI	Not applicable.
EXUS	EXUS EDAE will be a software output that will be part of the internal EXAI platform that will be used as a trade secret for the further development and enhancement of the EFS product offering.
FHG	The Privacy Risk module, its integration into TSM, TANGO PEP, and the showcase of its application in the IDIADA use case are relevant outputs.
FMAKE	At the moment, no other output is expected. If needed, we might print a physical object as final validation; but this is unknown at the moment.
FN	A roadmap to the sustainability of TANGO solution will be developed. It will include a sustainability plan, a value proposition definition, an analysis of possible business models, and the definition of a business plan for a sustainable vehicle to assume responsibility for ensuring the availability and maintenance of the key project results. We may also contribute to research outputs through white papers/scientific publications.
FSDE	The TSM module, its integration into the TANGO PEP, and the showcase of its application in the IDIADA use case are relevant outputs.
FUJ_LU	Not applicable.
IDIADA	The TSM module, its integration into the TANGO PEP, and the showcase of its application in the IDIADA use case are relevant outputs.
INTRA	INTRA, as system integrator, will not generate any research outputs.
KUL	Deliverables and other scientific publications, such as the TANGO White Paper.

6.1. Please identify research outputs other than data that you will generate during the TANGO project. Such outputs may be either digital (e.g., software, workflows, protocols, models, etc.) or physical (e.g., samples, new materials, etc.).		
LC	Digital research outputs such as branding kits, visual media (logos, infographics, videos), newsletters, social media content, and policy briefs. These outputs are shared publicly via the project website and trusted repositories. Documentation, metadata, and licensing ensure compliance with FAIR principles, promoting reuse and discoverability.	
MET	Not applicable.	
NORB	No other research output will be generated.	
QBE	Not applicable.	
SQD	**T2.2**	definition of RIAS functional and technical requirements (Models)
	T2.4	definition of TANGO Systems functional and technical requirements (Models)
	T2.5	definition of TANGO Security and privacy by design architecture (Models;Workflows)
	T3.3	definition of policy-based data access control (Models;Workflows)
	T3.4	delivery of the distributed self-encryption / decryption service (Service)
	T3.5	delivery of standardized levels of data handling security (Service)
	T5.4	definition of TANGO Privacy and threat modelling (Models)
	T5.4	delivery of PEC Privacy Enhancing component (Software)
	T6.4	delivery of Web based Data visualization user interfaces (Software)
	T6.3	delivery of Mobile and Web Front end user interfaces (Software)
UMU	Software for the technical solutions—e.g., generating, managing, and verifying p-ABCs, sticky policies, etc. Research publications on the topics UMU is involved. Participation in the definition of the workflows, models, and architectural design achieved in the project.	
UoG	Not applicable.	
UoM	A model trained to predict the energy efficient devices in a system.	
UPRC	Scientific publications.	
UTH	Software for device continuous behavioural authentication and the respective data model.	
VISAR	Not applicable.	
VTT	**T3.2**	An academic publication is considered.
	T4.1	Software.
XLAB	Not applicable.	

6.2. Please explain in sufficient detail how these research outputs will be managed, shared, and made available for reuse. In doing so, please consider the FAIR principles identified above.	
ABI LAB & NTT	The methodology developed is potentially valid for all institutions working in the retail banking sector; however the software components developed (the model and the data preparation pipelines) must be customized on a case-by-case basis based on the available data, the statistical distributions of the data, and the patterns of fraud suffered.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Project results like workflows or architecture are made public through public deliverables. Atos is the sole owner of Intellectual Property rights for the ABE toolset (solutions are not co-developed with other partners) and Federated Learning frameworks. In the case of FL, usage allowed for project partners under FAIR principles. Licensing to be agreed between ATOS and third parties.
CEA	Software is managed using git, and shared and stored using a GitLab instance private to CEA.
CESGA	Not applicable.
DBC	Not applicable.
DUT	The plugin will be made available online.
ECO	Not applicable.
EGI	Not applicable.
EXUS	The provided output will not be publicly available and will remain as a trade secret based also on the above response.
FHG	These outputs will be documented in related TANGO deliverables.
FMAKE	The data is publicly published, and a scientific paper has been written (part of a past project), explaining the data.
FN	All of the aforementioned information will be available on the project's shared space, in text form. Open access publication sites will be preferred in the case of publishing white papers/scientific publications. The TANGO project will be acknowledged everywhere.
FSDE	These outputs will be documented in related TANGO deliverables. The docker image of the TSM is available in the TANGO harbour environment.
FUJ_LU	Not applicable.
IDIADA	These outputs will be documented in related TANGO deliverables. The docker image of the TSM is available in the TANGO harbour environment.
INTRA	Not applicable.
KUL	Deliverables and other publications include keywords and abbreviations that facilitate navigating through them, in widely used formats (e.g., .pdf) and open access will be provided as appropriate.
LC	All research outputs—such as branding kits, visual media, newsletters, social media content, and policy briefs—are managed and shared in accordance with the FAIR (Findable, Accessible, Interoperable, Reusable) principles. These outputs are stored

6.2. Please explain in sufficient detail how these research outputs will be managed, shared, and made available for reuse. In doing so, please consider the FAIR principles identified above.		
	in standard, widely used formats (e.g., PDF, MP4, PNG) and are deposited in trusted open-access repositories such as Zenodo, where they are assigned persistent identifiers (e.g., DOIs) to support long-term findability. Accessibility is ensured by making these materials publicly available via HTTPS and through the TANGO project website. Reuse is promoted by applying standard open licenses (e.g., Creative Commons Attribution), and interoperability is supported by using commonly accepted formats and metadata standards. The outputs are documented with supporting materials (e.g., communication plans, branding guidelines, or readme files) that explain their context, intended use, and any technical considerations. These measures collectively facilitate broad and effective reuse by external stakeholders, including policymakers, researchers, and the wider public	
MET	Not applicable.	
NORB	Not applicable.	
QBE	Not applicable.	
SQD	Compliance with FAIR principles will be central to all activities, both technically as well as ethically in order to assure that TANGO results will be findable, accessible, interoperable and re-usable, with further details pending future definition (present status to be updated further during the project).	
UMU	Software will be published as open access—e.g., as code repositories (GIT). Research publications will be made open access (through gold open access or at least with manuscripts published in public repositories like ZENODO). Other project outputs like workflows or architecture are made public through public deliverables.	
UoG	Not applicable.	
UoM	The model will be usable through the TornadoVM API, in order to be inferred to make predictions.	
UPRC	Open access through Google Scholar or other repository.	
UTH	Software will be uploaded on GitHub repository, where appropriate documentation through README files will be provided.	
VISAR	Not applicable.	
VTT	**T3.2**	Open Access
	T4.1	Please provide your answer.
XLAB	Not applicable.	

Annex VIII – Resource Allocation

7.1. What costs will you incur for making data or other research outputs FAIR (e.g., direct and indirect costs related to storage, archiving, reuse, security, etc.)?	
ABI LAB & NTT	The data will not be accessible, so all management costs will be borne by the banks.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	The costs are described in the CEA indirect cost of the project.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Not applicable.
FMAKE	Hosting of the website.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	PAT requires updates to meet FAIR requirements. One of PAT's key characteristics is its secure and private design, which already covers the security aspect. However, for data manipulation, JSON data is somewhat FAIR within the scope of PAT.
IDIADA	Not applicable.
INTRA	Not applicable.
KUL	No specific/significant costs are envisaged.
LC	Costs include repository fees, secure storage, data curation, and communication team personnel effort.
MET	Not applicable.
NORB	The data storage capacity for the duration of the Pilot.
QBE	Not applicable.
SQD	Not yet defined.
UMU	No specific costs incurred.

7.1. What costs will you incur for making data or other research outputs FAIR (e.g., direct and indirect costs related to storage, archiving, reuse, security, etc.)?		
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	Not applicable.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made FAIR.
XLAB	Not applicable.	

7.2. How these costs will be covered?	
ABI LAB & NTT	The data will not be accessible, so all management costs will be borne by the banks.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	The costs are described in the CEA indirect cost of the project.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.
FHG	Not applicable.
FMAKE	Covered by project budgets; or indirectly by internal budgets.
FN	Not applicable.
FSDE	Not applicable.
FUJ_LU	The costs associated with making data or other research outputs FAIR can be covered through various means, including project budgets, collaborations, and partnerships with other organizations, and grants from industry-specific funding bodies.
IDIADA	Not applicable.

7.2. How these costs will be covered?	
INTRA	Not applicable.
KUL	If costs are incurred, these can be covered through the project funding, while acknowledging as appropriate the funding received by the EU.
LC	These costs are covered by the WP8 budget allocated to Lisbon Council or via internal resources.
MET	Not applicable.
NORB	Cost will be covered by the partner providing the cloud infrastructure for the relevant pilot.
QBE	Not applicable.
SQD	Not yet defined.
UMU	Through project's funding.
UoG	Not applicable.
UoM	Not applicable.
UPRC	Not applicable.
UTH	Within the project budget.
VISAR	Not applicable.
VTT	**T3.2** Not applicable.
	T4.1 Data will not be made FAIR.
XLAB	Not applicable.

7.3. How will long-term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept, and for how long).	
ABI LAB & NTT	The data will be managed by the banks.
ACV	Not applicable.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Long-term preservation is ensured thanks to both backups and dissemination of results in publication, and publication of pre-prints in open archives. As such, results are expected to be accessible for a long time.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.

7.3. How will long-term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept, and for how long).		
EXUS	Not applicable.	
FHG	Regular FHG retention measures apply.	
FMAKE	The target is a few years after the project has ended.	
FN	Not applicable.	
FSDE	Regular Fujitsu retention measures apply.	
FUJ_LU	For the duration of the project, no new necessary resources would be required. PAT is deployed and will be removed, along with its data, as soon as the project has ended.	
IDIADA	Regular IDIADA retention measures apply.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Long-term preservation follows Lisbon Council's data policy. Public materials are stored in trusted repositories and on the TANGO website for five years post-project. Internal backups are retained securely and deleted after five years.	
MET	The data will be stored on-premises at METRO until the end of the project.	
NORB	Data will be deleted after the Pilot.	
QBE	Not applicable.	
SQD	Not yet defined.	
UMU	Through the use of public, free of charge repositories.	
UoG	Not applicable.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	Two excessive months after the completion of the project.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	Data will not be made FAIR.
XLAB	Not applicable.	

Annex IX – Ethics

As explained in section 3, the project has gone through the discussion about the usage of personal data and decided to go for the signature of a Data Processing Agreement (DPA) per pilot, as well as the usage of consent forms for each user of the platform during the evaluation phase. The DPA has been discussed among all partners legal departments under the coordination of the legal partner (DBC) and the project coordinator (ATOS).

8.1. Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).	
ABI LAB & NTT	NTT Data does not access any data, so banks manage legal and ethical aspects according to the designated offices and functions.
ACV	No, besides the DPA.
AHOP	No.
ANYS	No, besides the DPA.
ATOS	No, besides the DPA.
CEA	No, besides the DPA.
CESGA	No, besides the DPA.
DBC	Not applicable.
DUT	No, besides the DPA.
ECO	Not applicable.
EGI	Not applicable.
EXUS	No, besides the DPA.
FHG	No, besides the DPA.
FMAKE	No, besides the DPA.
FN	Not applicable.
FSDE	No, besides the DPA.
FUJ_LU	No, besides the DPA.
IDIADA	No, besides the DPA.
INTRA	No, besides the DPA.
KUL	No.
LC	Yes. Personal data collected (e.g., event participants, newsletter subscribers) are subject to GDPR. Personal data are not shared publicly. Ethics compliance is documented in the project's Description of the Action (DoA).
MET	No, besides the DPA.
NORB	No, there won't be any legal issues.

8.1. Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).	
QBE	Sharing of biometric data could be an issue. Covered by the DPA, consent forms and deletion of data after the evaluation of the pilots.
SQD	N/A—SQD will not use its own proprietary datasets. Datasets from RIAS are still under definition but personal data collection, transformation, and sharing will respect Portuguese government issued legislation regarding GPDR principles. SQD must inform that collection of data such as racial/ethnical/political/religion/association with unions, and others belonging to individuals are specifically forbidden under Portuguese Law; therefore those information classes are not inquired or collected and will not be part of any TANGO dataset. Please refer to https://dre.pt/dre/en/detail/act/58-2019-123815982.
UMU	No, besides the DPA.
UoG	No, besides the DPA.
UoM	No, besides the DPA.
UPRC	No.
UTH	Data is not intended to be shared.
VISAR	Not applicable.
VTT	**T3.2** No
	T4.1 No, besides the DPA.
XLAB	Not applicable.

8.2. Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?	
ABI LAB & NTT	No, data sharing is foreseen.
ACV	Yes.
AHOP	Not applicable.
ANYS	Not applicable.
ATOS	Not applicable.
CEA	Not applicable.
CESGA	Not applicable.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	Not applicable.

8.2. Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?		
FHG	Yes.	
FMAKE	No personal data included in the dataset.	
FN	Not applicable.	
FSDE	Yes.	
FUJ_LU	Yes.	
IDIADA	Yes.	
INTRA	Not applicable.	
KUL	Not applicable.	
LC	Yes. Consent forms include clear information about data sharing, storage, and long-term preservation.	
MET	No personal data will be shared.	
NORB	Yes.	
QBE	The app will inform the user, and informed consent will be gathered for the participants of the evaluation.	
SQD	Yes, as regulated per Portuguese GPDR legislation https://dre.pt/dre/en/detail/act/58-2019-123815982.	
UMU	Not applicable.	
UoG	Not applicable. However, if there is need to collect personalised data, consent and data handling measures will be included in the questionnaires.	
UoM	Not applicable.	
UPRC	Not applicable.	
UTH	Consent for monitoring of data will be provided by the users involved in the pilot evaluation.	
VISAR	Not applicable.	
VTT	**T3.2**	Not applicable.
	T4.1	No.
XLAB	Not applicable.	

Annex X – Other Issues

8.3. Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?	
ABI LAB & NTT	No.
ACV	No.
AHOP	No.
ANYS	No.
ATOS	No.
CEA	No.
CESGA	No.
DBC	Not applicable.
DUT	Not applicable.
ECO	Not applicable.
EGI	Not applicable.
EXUS	EXUS complies with all provisions and articles of General Data Protection Regulation (GDPR) and data management procedures under the ISO: 27001 standards.
FHG	No.
FMAKE	No.
FN	No.
FSDE	No.
FUJ_LU	No.
IDIADA	No.
INTRA	No.
KUL	No.
LC	Yes. LC follows internal data protection and IT security policies, Horizon Europe guidelines, and GDPR requirements.
MET	No.
NORB	No.
QBE	No.
SQD	Not applicable—SQD will not use its own proprietary datasets. Datasets from RIAS are still under definition but personal data collection, transformation, and sharing will respect Portuguese government issued legislation regarding GPDR principles. SQD must inform that collection of data such as racial/ethnical/political/religion/association with unions, and others belonging to individuals are specifically forbidden under Portuguese Law; therefore, those

8.3. Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?		
	information classes are not inquired or collected and will not be part of any TANGO dataset. Please refer to https://dre.pt/dre/en/detail/act/58-2019-123815982.	
UMU	No.	
UoG	No.	
UoM	No.	
UPRC	Not applicable.	
UTH	No.	
VISAR	No.	
VTT	**T3.2**	Not applicable.
	T4.1	No.
XLAB	No.	

Annex XI – EXUS dataset

Pilot	Field Name	Field Type
Smart hospitality	Timestamp	date
	refSensor	Sensor name
	temperature	numeric
	humidity	numeric
	pressure	numeric
	battery	numeric
	link quality	numeric
	illuminance	numeric
	presence	numeric
	occupancy	numeric
Smart manufacturing	Frame	image
	Laser params (Speed, power	numeric
	Position (x, y)	numeric
	Scan line index	numeric
Smart Retail	Product Code (Unique Key)	string
	Product Description	string
	Product's Main Supplier Code	string
	Product's Category Hierarchy 1st Level Code	string
	Product's Category Hierarchy 2nd Level Code	string
	Product's Category Hierarchy 2nd Level Description	string
	Product's Category Hierarchy 3rd Level Code:	string
	Product's Category Hierarchy 3rd Level Description	string
	Product Entry Date	string
	Product's Measurement Unit Code	string

Pilot	Field Name	Field Type
	Product's Measurement Unit Description	string
	Product's Active Indicator	numeric
	Master Product Code for Product Families (reference ProductCode)	string
	Product's Life Cycle Status Code	numeric
	Product's Life Cycle Status Description	string
	Product's Buyer Code	string
	Product's Buyer Description	string
	Product's Buyer Manager Code	string
	Product's Buyer Manager Description	string
	Customer Code (Unique Key)	string
	Master Customer Code for Customer Groups (Reference Customer Code)	string
	Customer Inactive Indicator	numeric
	Customer's Entry Date	string
	Customer's Profession Group Code	string
	Customer's Profession Group Description	string
	Customer's Profession Code	string
	Customer's Profession Description	string
	Company Description (Unique Key)	string
	Transaction Date (Unique Key)	string
	Transaction Time	string
	Customer Code or Shop Code for retail transactions (reference TangoCus_sample)	string
	Transaction Id (Unique Key)	numeric
	Shop Code	string
	Retail Sale Transaction Indicator	numeric

Pilot	Field Name	Field Type
	Item Code (reference TangoProd_sample)	string
	Sale Quantity	numeric
	Sale Value including VAT	numeric

Annex XII – NORB dataset

INDIVIDUAL

Field Name	Field Type	Semantics
PDTLS_Gre	single_select	Greeting
PDTLS_Fname	text	First Name of an individual
PDTLS_Lname	text	Last Name of an individual
PDTLS_MaiName	text	Maiden Name of an individual
PDTLS_MidName	text	Middle Name of an individual
PDTLS_IDType	single_select	List of possible ID types (e.g., National ID, Passport)
PDTLS_IDNo	text	National ID Number
PDTLS_IDCountry	single_select	National ID Country
PDTLS_IDIssDat	date	National ID Issue Date
PDTLS_IDExpDat	date	National ID Expiry Date
PDTLS_PassNo	text	Passport Number
PDTLS_PassCountry	single_select	Passport Country
PDTLS_PassIssDat	date	Passport Issue Date
PDTLS_PassExpDat	date	Passport Expiry Date
PDTLS_DOB	date_tz_independent	Date of Birth of an individual
PDTLS_BirthCountry	single_select	Country of Birth
PDTLS_BirthCity	text	City of Birth
PDTLS_Gend	single_select	Gender
PDTLS_PermResCountry	single_select	Country of Permanent Residence
PDTLS_ResStat	text	Residence Status in host country (e.g., Citizenship, Green Card Holder, non-resident, etc.)
PDTLS_ResVisa_Type	single_select	Type of VISA in case residency status is VISA holder

Field Name	Field Type	Semantics
PDTLS_ResVisa_GCardNo	text	VISA/Green Card Number
PDTLS_ResVisa_GcardIssDat	date	VISA/Green Card Issue Date
PDTLS_ResVisa_GCardExpDat	date	VISA/Green Card Expiry Date
PDTLS_ATTACH_IDPassScan	file	National ID or Passport scan/image file
PDTLS_ESignature	e_signature	E Signature object
PDTLS_MarStat	single_select	Marital Status
PDTLS_MultCitiz	boolean	Multiple Citizenship holder
PDTLS_MultCitizCountry	multi_select	Other Citizenship Countries
PDTLS_ATTACH_MultCitizIDProof	file	National ID or Passport scan/image file
PDTLS_Lang	multi_select	Mother tongue and other languages known
PDTLS_OtherNames	boolean	Ever had or been known by any other names
PDTLS_OtherNamesFName	text	Other First Name of an individual
PDTLS_OtherNamesMName	text	Other Middle Name of an individual
PDTLS_OtherNamesLName	text	Other Last Name of an individual
EMPDTLS_EmpType	single_select	Employment Type
EMPDTLS_EmpTypeOther	single_select	Other Employment Type
EMPDTLS_EmpCat	multi_select	Sector of Employment based on NACE codes and description
EMPDTLS_Occu	text	Occupation/Role-Title
EMPDTLS_Employer	text	Employer
EMPDTLS_EmpStartDate	date	Employment Start Date
EMPDTLS_ATTACH_OccuProof	file	Proof of occupation (payslip etc.) scan/image file
EMPDTLS_Share	numeric	Shareholder %
EMPDTLS_Countries	multi_select	Countries where the business operates

Field Name	Field Type	Semantics
RESADD_Add1 - Street	text	Residential Address Address Line 1 cfc:AddressFix_Type - Street
RESADD_Add2 - Building/No	text	Residential Address Address Line 2 cfc:AddressFix_Type - BuildingIdentifier
RESADD_Add3 - District	text	Residential Address Address Line 3 cfc:AddressFix_Type DistrictName
RESADD_Country	single_select	Residential Address Country
RESADD_City	text	Residential Address City cfc:AddressFix_Type - City
RESADD_PostCod	text	Residential Address Postal Code cfc:AddressFix_Type - PostCode
RESADD_ATTACH_UtilityBill	file	Proof of address (e.g., past 6 month utility bill scan/image) file
RESADD_ResType	single_select	Type of residence
CONDTLS_HomePNo	text	Home Phone Number
CONDTLS_BusPNo	text	Business Phone Number
CONDTLS_Mob	text	Mobile Phone Number
CONDTLS_Fax	text	Fax Number
CONDTLS_Email	text	Email
FATCACRS_HCCitizen	boolean	Citizen/resident of host country for tax purposes
FATCACRS_HCTIN	text	Host Country Taxpayer Identification Number
FATCACRS_USCitizen	boolean	Citizen/resident of US for tax purposes
FATCACRS_USTIN	text	US Taxpayer Identification Number
FATCACRS_OthCitizen	boolean	Citizen/resident of any other country (apart from host and US) for tax purposes
PEP_Flag	boolean	Relationship exists with a person who holds/held a prominent public function

Field Name	Field Type	Semantics
PEP_Relationship	text	Relationship with prominent public function person
PEP_Position	text	Prominent public function position
PEP_Country	single_select	Country
PEP_FromDt	date_tz_independent	Period in which prominent public function position is/was held From
PEP_ToDt	date_tz_independent	Period in which prominent public function position is/was held To
FININFO_TxnNature	text	Purpose of Opening Account - Nature of Transactions
FININFO_AccOpeAmt	numeric	Account initial opening amount
FININFO_AccCCY	single_select	Main Account Currency
FININFO_AnnCRTurn	numeric	Expected Annual Credit Turnover amount
FININFO_AnnCRTurnCcy	single_select	Expected Annual Credit Turnover amount Currency
FININFO_AnnCashTurn	numeric	Expected Annual Cash Turnover amount
FININFO_AnnCashTurnCcy	single_select	Expected Annual Cash Turnover amount Currency
FININFO_AnnDRTurn	numeric	Expected Annual Debit Turnover amount
FININFO_AnnDRTurnCcy	single_select	Expected Annual Debit Turnover amount Currency
FININFO_AnnIncome	numeric	Annual Income amount
FININFO_AnnIncomeCcy	single_select	Annual Income amount Currency
FININFO_IncomeSrc	multi_select	Source of Income
FININFO_OpenAmtSrc	multi_select	Opening amount source of funds
FININFO_OpenAmtTrfr	multi_select	Means of transferring the opening amount
FININFO_OthBankRel	boolean	Relationship with another Bank

Field Name	Field Type	Semantics
FININFO_OthBankName	text	Other Bank name
FININFO_NetWorth	numeric	Net worth amount
FININFO_NetWorthCCY	single_select	Net worth currency
FININFO_Assets	text	Assets available
FININFO_AssetsSrc	text	Source of Assets
FININFOEDD_MontCashPerc	numeric	Percentage of cash transactions on a monthly basis
FININFOEDD_CashMontNo	numeric	Expected total cash transactions per month
FININFOEDD_CashMontAvg	numeric	Average cash transaction amount
FININFOEDD_CashMontAvgCCY	single_select	Average cash transaction currency
FININFOEDD_CashMontTot	numeric	Total monthly cash transaction amount
FININFOEDD_CashMontTotCCY	single_select	Total monthly cash transaction currency
FININFOEDD_ MontChqPerc	numeric	Percentage of cheque transactions on a monthly basis
FININFOEDD_ChqMontNo	numeric	Expected total cheque transactions per month
FININFOEDD_ChqMontAvg	numeric	Average cheque transaction amount
FININFOEDD_ChqMontAvgCCY	single_select	Average cheque transaction currency
FININFOEDD_ChqMontTot	numeric	Total monthly cheque transaction amount
FININFOEDD_ChqMontTotCCY	single_select	Total monthly cheque transaction currency
FININFOEDD_MontWirePerc	numeric	Percentage of wire transfer transactions on a monthly basis
FININFOEDD_WireMontNo	numeric	Expected total wire transactions per month
FININFOEDD_WireMontAvg	numeric	Average wire transaction amount
FININFOEDD_WireMontAvgCCY	single_select	Average wire transaction currency

Field Name	Field Type	Semantics
FININFOEDD_WireMontTot	numeric	Total monthly wire transaction amount
FININFOEDD_WireMontTotCCY	single_select	Total monthly wire transaction currency
FININFOEDD_MajRegRemitt	text	Major/Frequent Remitters
FININFOEDD_MajRegRemittOrigin	multi_select	Major/Frequent Remitters country
FININFOEDD_MajRegBene	text	Major/Frequent Beneficiaries
FININFOEDD_MajRegBeneDest	multi_select	Major/Frequent Beneficiaries country

CORPORATE

fieldName	fieldtype	semantics
CorCD_CompName	text	Name of the Company
CorCD_RegNum	text	Company's Registration Number
CorCD_RegCountry	single_select	Country of Registration/Incorporation
CorCD_RegDate	date_tz_independent	Registration Date
CorCD_USEIN	text	Federal Employer Identification Number (FEIN) or the Federal Tax Identification Number
CorCD_LEINum	text	Legal Entity ID
CorCD_TypeLE	single_select	Type of Business Structure - Legal Entity Type - Specific to each host country
CorCD_TypeLEOther	text	Legal Entity Type Other
CorCD_BusActivitSect	single_select	Business activity sector
CorCD_MainBusActivit	text	Main business activity - Description of business nature and operating activities
CorCD_SecBusActivit	text	Secondary business activity
CorCD_RequiredLicenses	text	Licenses required for company's operation
CorCD_RequiredLicensesProof	file	Proof of required licenses

fieldName	fieldtype	semantics
CorCD_IsPubliclyList	boolean	Publicly listed company
CorCD_StockExchange	single_select	Stock Exchange Market where the company is listed
CorCD_StockTrdName	text	Trading name in stock exchange
CorCD_ISIN	text	ISIN number
CorCD_BearerShares	boolean	Bearer shares available
CorCD_CompanyCert	file	Company certificate/Certificate of Incorporation (or foreign equivalent)
CorCD_OrgChart	file	Organizational Chart showing all entities between the company and ultimate parent company, clearly showing the ownership and %. The chart must include full legal name, address, registration number, and country of incorporation for parent company/ultimate parent company, or be supplemented with official document providing the same information
CorCD_NumEmp	numeric	Number of employees in the company
CorCD_Website	text	Company's Website
CorCD_ComUnderReg	boolean	Registered company
CorCD_FormationDoc	file	Company's Formation document(s)
CorCD_ParentCoRegCountry	single_select	Parent Company's Country of incorporation
CorCD_RegCert	file	Parent Company's Registration Certificate
CorCD_InternationalBus	multi_select	Countries where the company has international business. When high risk countries are selected, EDD is triggered. High risk countries (Cuba, Iran, NORBth Korea, Sudan, Syria, Crimea)
CorCD_Crypto	boolean	Is your company involved with mining or brokering of crypto currencies or does your company accept crypto currencies as payment?

fieldName	fieldtype	semantics
CorCD_Diamonds	boolean	Is your company involved with rough diamonds or gold trade?
CorCD_Suppliers	text	Type of suppliers
CorCD_PrimaryMarket	multi_select	Countries that are primary market and primary source of revenue
CorCD_INV	boolean	Is the entity an investment entity? (e.g., trading in markets instruments, P.E. fund, fund management, a broker etc.)
CorCD_GovRel	boolean	Relationship with governmental authorities and/or agencies statutory corporations and/or registered public companies (PLC)
CorCD_GovRelDetails	text	Relationship type with governmental authorities and/or agencies statutory corporations and/or registered public companies (PLC) and bodies involved
CorCD_GovRelShares	numeric	Governmental shareholding
CorContD_ContactPersonFName	text	Primary contact person first name
CorContD_ContactPersonLName	text	Primary contact person last name
CorContD_ContactPersonMidName	text	Primary contact person middle name
CorContD_ContPerPosition	text	Position role in the company
CorContD_ContPerTel	text	Primary contact person Telephone
CorContD_ContPerEmail	text	Primary contact person Email
CorAddr_RegAddrAdd1Str	text	Registered Address: Address Line 1 cfc:AddressFix_Type - Street
CorAddr_RegAddrCountry	single_select	Registered Address Country
CorAddr_RegAddrCity	text	Registered Address City
CorAddr_RegAddrPostCod	text	Registered Address Postal Code
CorAddr_RegAddrPOB	text	Registered Address P.O. Box
CorIndv_Gre	single_select	Greeting

fieldName	fieldtype	semantics
CorIndv_Fname	text	First Name
CorIndv_LName	text	Last Name
CorIndv_MidName	text	Middle Name
CorIndv_FatherName	text	Father's Name
CorIndv_MaiName	text	Maiden Name
CorIndv_Gend	single_select	Gender
CorIndv_IDType	single_select	List of possible ID types (e.g., National ID, Passport)
CorIndv_IDNo	text	National ID Number
CorIndv_IDCountry	single_select	National ID Country
CorIndv_IDIssDat	date	National ID Issue Date
CorIndv_IDExpDat	date	National ID Expiry Date
CorIndv_PassNo	text	Passport Number
CorIndv_PassCountry	single_select	Passport Country
CorIndv_PassIssDat	date	Passport Issue Date
CorIndv_PassExpDat	date	Passport Expiry Date
CorIndv_PasspIDCopy	file	Copy of passport
CorIndv_Nationality	single_select	Nationality
CorIndv_DOB	date_tz_independent	Date of Birth
CorIndv_BirthCountry	single_select	Country of Birth
CorIndv_BirthCity	text	City of Birth
CorIndv_Signature	e_signature	E Signature object
CorIndv_Add1Str	text	Address Line 1 cfc:AddressFix_Type - Street
CorIndv_Add2BuildNo	text	Address Line 2 cfc:AddressFix_Type - BuildingIdentifier

fieldName	fieldtype	semantics
CorIndv_Add3Distr	text	Address Line 3 cfc:AddressFix_Type DistrictName
CorIndv_Country	single_select	Address Country
CorIndv_City	text	Address City
CorIndv_PostCod	text	Address Postal Code
CorIndv_POB	text	Address P.O. Box
CorIndv_Tel	text	Director's Telephone
CorIndv_Email	text	Director's Email
CorIndv_PEP	single_select	Director considered as a politically exposed person - PEP or related person to PEP
CorIndv_PEPRelationship	text	Relationship with prominent public function person
CorIndv_PEPPosition	text	Prominent public function position
CorIndv_PEPCountry	single_select	Country
CorIndv_PEPFromDt	date_tz_independent	Period in which prominent public function position is/was held From
CorIndv_PEPToDt	date_tz_independent	Period in which prominent public function position is/was held To
CorIndv_Dirct	boolean	Individual is a director in the company
CorIndv_Shhld	boolean	Individual is a shareholder in the company
CorIndv_Signatory	boolean	Individual is an authorized representative of the company
CorIndv_UBO	boolean	Individual is an ultimate beneficial owner
CorIndv_ShhldShare	numeric	Shareholder's Share % - when individual's authority is shareholder
CorShhld_Flag	boolean	Shareholders available
CorShhld_Classes	boolean	Shares divided into several classes, where one or several hold control beyond

fieldName	fieldtype	semantics
		number of shares? If yes, attach explanation and shareholder book
CorShhld_Book	file	Shareholder Book and shares classes explanation
CorShhld_Type	boolean	Shareholding LE
CorShhldCD_CompName	text	Name of the Company
CorShhldCD_RegNum	text	Company's Registration Number
CorShhldCD_RegCountry	single_select	Country of Registration/Incorporation
CorShhldCD_RegDate	date_tz_independent	Registration Date
CorShhldCD_USEIN	text	Federal Employer Identification Number (FEIN) or the Federal Tax Identification Number
CorShhldCD_LEINum	text	Legal Entity ID
CorShhldCD_TypeLE	single_select	Type of Business Structure - Legal Entity Type - Specific to each host country
CorShhldCD_TypeLEOther	text	Legal Entity Type Other
CorShhldCD_BusActivitSect	single_select	Business activity sector
CorShhldCD_MainBusActivit	text	Main business activity - Description of business nature and operating activities
CorShhldCD_SecBusActivit	text	Secondary business activity
CorShhldCD_RequiredLicenses	text	Licenses required for company's operation
CorShhldCD_RequiredLicensesProof	file	Proof of required licenses
CorShhldCD_IsPubliclyList	boolean	Publicly listed company
CorShhldCD_StockExchange	single_select	Stock Exchange Market where the company is listed
CorShhldCD_StockTrdName	text	Trading name in stock exchange
CorShhldCD_ISIN	text	ISIN number
CorShhldCD_BearerShares	boolean	Bearer shares available

fieldName	fieldtype	semantics
CorShhldCD_CompanyCert	file	Company certificate/Certificate of Incorporation (or foreign equivalent)
CorShhldCD_OrgChart	file	Organizational Chart showing all entities between the company and ultimate parent company, clearly showing the ownership and %. The chart must include full legal name, address, registration number, and country of incorporation for parent company/ultimate parent company, or be supplemented with official document providing the same information
CorShhldCD_NumEmp	numeric	Number of employees in the company
CorShhldCD_Website	text	Company's Website
CorShhldCD_ComUnderReg	boolean	Registered company
CorShhldCD_FormationDoc	file	Company's Formation document(s)
CorShhldCD_ParentCoRegCountry	single_select	Parent Company's Country of incorporation
CorShhldCD_RegCert	file	Parent Company's Registration Certificate
CorShhldCD_InternationalBus	multi_select	Countries where the company has international business. When high risk countries are selected, EDD is triggered. High risk countries (Cuba, Iran, NORBth Korea, Sudan, Syria, Crimea)
CorShhldCD_Crypto	boolean	Is your company involved with mining or brokering of crypto currencies or does your company accept crypto currencies as payment?
CorShhldCD_Diamonds	boolean	Is your company involved with rough diamonds or gold trade?
CorShhldCD_Suppliers	text	Type of suppliers
CorUBO_Occupation	text	Beneficial Owner's Occupation - when individual's authority is UBO

fieldName	fieldtype	semantics
CorUBO_ResAdd1Str	text	Residence Address Address Line 1 cfc:AddressFix_Type – Street - when individual's authority is UBO
CorUBO_ResAdd2BuildNo	text	Residence Address: Address Line 2 cfc:AddressFix_Type - BuildingIdentifier- when individual's authority is UBO
CorUBO_ResAdd3Distr	text	Residence Address: Address Line 3 cfc:AddressFix_Type DistrictName- when individual's authority is UBO
CorUBO_ResCountry	single_select	Residence Address Country - when individual's authority is UBO
CorUBO_ResCity	text	Residence Address City - when individual's authority is UBO
CorUBO_ResPostCod	text	Residence Address Postal Code - when individual's authority is UBO
CorUBO_ResPOB	text	Residence Address P.O Box - when individual's authority is UBO
CorUBO_FATCACRS_HCCitizen	boolean	Citizen/resident of host country for tax purposes - when individual's authority is UBO

Annex XIII – Notice on Personal Data Processing and Providing of Consent Form

As part of the TANGO project, we are conducting a pilot at [...], a [...]. This pilot aims to […]

The Pilot will take place from __/__/2025 until __/__/2025.

We are committed to protecting your privacy and handling your personal data responsibly, ensuring compliance with the General Data Protection Regulation (GDPR).

Before commencing your participation, we kindly ask you to read the following information and provide your written consent by signing below.

Please note that we value your free and voluntary participation in this pilot in order to ensure that you do not feel compelled to participate. It is important for us to understand your experience during this pilot. Thus, we will ask you to complete a brief questionnaire before, during, and after the pilot.

1. Data Controller

The Data Controller will be (…), as part of the TANGO Project, as well as the TANGO Project Consortium.

The Data Protection Officer will be the designated legal partner of TANGO Project, "Diadikasia Business Consulting Symvouloi Epicheiriseon AE"; you can reach out to them via an email, should you have any questions regarding your personal data, in the following email address: legalint@diadikasia.gr.

2. Purpose of Data Collection

[In this section, the specific action of the pilot (task, workshop, interview, seminar, questionnaire etc.) should be described, along with the purpose of the personal data processing. For example:

1. *The responses of this questionnaire will be crucial in enhancing workplace safety and protection of industrial property by identifying when employees enter certain sensitive areas.*
2. *The purpose of the processing operations of this action is the […] of the TANGO project.*]

3. Data Collected

[In this action the categories of personal data being processed should be mentioned in detail, as well as all the processing operations of the personal data. It is important that, depending on the specific processing operation, the necessary information is provided to ensure full transparency and obtain the valid consent of the participant. Additionally, the purpose of the data processing must be stated to ensure that we comply with the principle of data minimisation, as defined in the GDPR (only the data deemed necessary for the purpose of the specific action is collected, and no more).

For example:

We will be processing the following personal data:

i. *Name and surname*
ii. *Place of occupation*
iii. *Work title*
iv. *Signature*
v. *Email address*

The envisaged processing operations on the aforementioned personal data are the following:

vi. *Collection directly from you*

vii. *Storage on […] until […]*

viii. *Use in order to enhance workplace safety and protection of industrial property*

ix. *Use in order to coordinate the procedure of the interviews, communicate with you and proceed to the relevant arrangements*]

4. How Will Your Data Be Used?

[In this section, the specific use of the collected data should be described. For example:

1. *Your data will be used to assess safety risks, optimise workflows, and evaluate the performance of the TANGO technologies.*
2. *Data will be aggregated and anonymised for reporting purposes and will not be used for purposes other than those mentioned here.*]

5. Legal Basis for Processing

Your personal data will be processed according to the relevant legal provisions. The legal basis for the processing of your personal data during the pilot is your consent relating to Article 6 paragraph 1 sentence 1a of the General Data Protection Regulation.

Please note that you have the right to withdraw your consent at any time, without any negative consequences. If you choose to withdraw consent, this will not affect the lawfulness of any data processing that has already taken place before your withdrawal. You can contact legalint@diadikasia.gr to revoke consent.

6. Data controllers - Access to Personal Data

According to GDPR, 'controller' means the natural or legal person, public authority, agency, or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law. […] is the data controller, therefore the one that coordinates and processes the data collected.

Access to your data will be strictly limited to:

- Authorised personnel within the TANGO project, including data controller's (…) technical team.
- Third-party service providers (such as cloud service providers), who are GDPR-compliant.
- Other project partners may review anonymised data for research purposes.

7. Retention Period

Your personal data will be retained for as long as the Pilot is taking place; after the pilot's completion, it will then be either anonymised or securely deleted.

8. Your Rights

As a result of the data processing operations, according to the General Data Protection Regulation (GDPR), the data subject participant to the pilot can exercise at any time the following rights, when these apply:

i. **"Right to be informed,"** regarding anything that is related to the processing of personal data;

ii. **"Right of access"** to the personal data and any other information related to the data processing activities;

iii. **"Right to rectification"** of inaccurate personal data or completion of incomplete personal data;

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;

v. **"Right to restriction of processing"** of the personal data under certain conditions;

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;

vii. **"Right to object"** at any time to processing of personal data concerning him/her;

viii. **"Right to object to the automated individual decision-making, including profiling"**;

ix. **"Right of filling a complaint"** to the competent national supervisory authority;

x. **"Withdrawal from the provided consent"** at any time freely.

In case you wish to exercise your forementioned rights, please contact the designated legal partner of TANGO project, "Diadikasia Business Consulting Symvouloi Epicheiriseon AE," via an email in the following email address: legalint@diadikasia.gr.

Participant's statement

I declare that:

- I have read the notes written above and understand what the project and the pilot I am participating in are related to – in this case, the testing and implementation of the TANGO framework. I have been given the opportunity to ask questions and have had them answered to my satisfaction.
- I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason immediately, without consequences.
- I agree that my non-personal research data may be used by others for future research and/or reports, scientific publications or other dissemination and communication outputs. I am assured that the confidentiality of my personal data will be upheld through the removal of identifiers (anonymisation).
- I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the EU General Data Protection Regulation (Regulation EU 2016/679).

To this end, I also explicitly declare that:

I. ☐ I agree to participate in the pilot of the TANGO project

II. ☐ I provide my consent to the processing of my personal data for the purposes related to my participation in the TANGO pilot.

Signature: _____

Date: _____

Name: _____

Annex XIV – TANGO Pilots Questionnaire

The following questionnaire shall be completed by the partners participating in the pilot activities of the TANGO project (WP7 related activities). It serves as a monitoring tool for partners KUL and DBC to collect the necessary information related to the pilot activities, as a result of their respective duties and most crucially tasks 1.4 and 2.6 of the project. The aim of this joint action is the collection of as much information as possible in order to identify any possible risks or gaps and overall ensure the privacy, legal, and ethical compliance of the project related activities.

We would like to ask from the partners to provide us with as much detailed information as possible in order for us to be able to proceed to our monitoring activities effectively. The information received through this tool will be presented in reports and related deliverables (e.g., D2.4). Partners KUL and DBC remain at your disposal for any additional information or guidance in relation to the current questionnaire.

DATA MANAGEMENT

Personal data

Q	Is the data personal or non-personal, or is it a mixed dataset?
A	

Q	If there are any personal data, please specify which type of data that is (for example, name, surname, height, biometric data, etc.)
A	

Q	If there are any personal data, please specify whether they belong to one or more of these categories: genetic data, biometric data for the purpose of uniquely identifying a natural person, health data, data concerning a natural person's sex life or sexual orientation, data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, data concerning children or vulnerable adults.
A	

Q	What is the purpose of the processing of the aforementioned personal data?
A	

Q	What are the planned/foreseen actions (processing operations) related to the aforementioned personal data?
A	

Q	Will the processing operations be based in the provision of consent from the participants?
A	

Q	If no, on which legal basis will the processing operations be based?
A	

Q	Are the participants informed about their data protection related rights?
A	

Q	What security measures do you implement as an organisation that could be relatable to the project related activities?
A	

Q	Have you appointed a Data Protection Officer?
A	

Q	Is there any non-personal data that may be protected under EU, national, or contract law? E.g. trade secrets, statistical confidentiality.
A	

Data storage

Q	In which format is the data stored?
A	

Q	In which physical and virtual location is the data stored?
A	

Q	If you store personal data (not data overall) in your premises, please provide us with a reason for such processing operation.
A	

Q	Which is the foreseen period of storage for the data?
A	

Q	Will the data be transferred or stored to countries outside the EU/EEA?
A	

Q	Please explain whether any manipulations of the data will take place during the project (e.g. anonymisation, pseudonymisation).
A	

Q	Are there any measures taken to ensure data integrity, data quality, and security of the data? E.g. any backups that ensure data integrity.
A	

Data availability and sharing

Q	Is the data available to other TANGO partners?

A	

Q	For data that is available to other TANGO partners, please provide information about conditions for or restrictions to access to or use of the dataset (including, but not limited, to control of access to personal data/confidential information).
A	

Q	If data will be closed (or can be shared under restrictions), please specify which restrictions apply and provide an explanation.
A	

Q	Please indicate the expected time that data will be made open.
A	

Making data FAIR (findable, accessible, interoperable, and reusable)

Q	Will data be identified by a persistent identifier?
A	

Q	Will rich metadata be provided to allow discovery? If yes, please specify which metadata will be created and if search keywords will be provided in the metadata.
A	

Q	Does the data clearly and explicitly include the identifier of the data it describes?
A	

Q	Is the metadata registered or indexed in a searchable resource?

A	

Q	Is the metadata retrievable by their identifier using a standardised communications protocol that is open, free, and universally implementable?
A	

Q	Does the (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation?
A	

Q	Does the (meta)data use vocabularies that follow FAIR principles – meaning that the vocabulary used in describing the dataset shall be documented and resolvable using unique and persistent identifiers, with the information being easy to find and accessible by those who use the dataset?
A	

Q	Does the (meta)data include qualified references (i.e., cross-references explaining their intent) to other (meta)data?
A	

Q	Is the (meta)data richly described with a plurality of accurate and relevant attributes? This means that adequate and appropriate labels attached to the data should allow the (re)-user (human or machine) to decide if the data will indeed be useful for them in the particular context (by considering, for example, the context under which the data was originally generated).
A	

Q	Is the (meta)data released with a clear and accessible data usage license, clearly indicating the conditions under which the data can be reused?
A	

Q	Is the (meta)data associated with detailed provenance? i.e., it should be clear where the data comes from, who generated or collected it, how it has been processed, whether it has been published before, and whether it contains data from other sources.
A	

Q	Is the (meta)data associated with detailed provenance?
A	

Q	Does the (meta)data meet domain-relevant community standards?
A	

ETHICS

Human participation

Q	Will humans be involved as participants in the pilot? (third parties, not related to the project partners)
A	

Q	Are there any expected benefits to participants?
A	

Q	Will participants receive verbal or written information about the TANGO project? At what form will the participants receive the information?
A	

Q	Will information be provided in languages other than English?

A	

Q	Is the participation in the pilot of the TANGO project voluntary?
A	

Q	Is the participation in the TANGO pilot based on prior consent of the participant in terms of participation?
A	

Q	If yes, will the consent be provided verbally or on a written form (form etc.)?
A	

Q	If yes, is the participant informed that he/she can effortlessly withdraw from the provided consent?
A	

Q	Will your designated TANGO pilot involve vulnerable people? (e.g. children, people with learning disabilities, patients)
A	

Q	Will participants be reimbursed for any expenses incurred?
A	

Q	Do you implement any criteria during the selection of the participants? If yes, please explain the reason.
A	

Q	Have you ensured that no participant is excluded on the grounds of sexual orientation, age, gender, religious belief, ethnic group, or disability?
A	

Q	Will the activities involve prolonged or repetitive testing of the participants?
A	

Q	Is it possible, as a result of the project related activities, for psychological stress, anxiety, cause of harm, or any other negative consequences beyond the risks encountered in NORBmal life?
A	

Dissemination

Q	Is there an intention to publish or disseminate the outcomes of the TANGO related pilots that you lead/participate in? If yes, please explain the procedure to be followed.
A	

Q	Will the participants in the pilots receive feedback?
A	

Q	Will any reports/feedback include recommendations for improvement?
A	

Q	Will results be presented in a way that does not identify individuals? If no, please justify such action and refer to the measures that will be implemented.
A	

We would like to thank you for your cooperation. In case we identify any gaps, issues or in general inconsistencies, we will contact you in order to receive any additional information and ensure that the procedures are being followed in respect to the applicable legal, privacy, and ethical related requirements.