TRUSTEE D5.1 Report on TRUSTEE Pilot Campaign Plan and Evaluation Methodology

Work package	WP5: TRUSTEE MULTI-DATA SPACES Pilot Campaign
Task	Task 5.1: TRUSTEE Pilot Campaign Plan and Evaluation Methodology Task 5.2: Pilot use case 1: ENERGY dataset Task 5.3: Pilot use case 2: HEALTH Dataset Task 5.4: Pilot use case 3: EDUCATION Datasets Task 5.5: Pilot use case 4: AUTOMOTIVE Datasets Task 5.7: Pilot use case 6: Trusted multi- disciplinary data exchange
Authors	**Dimitra Papatsaroucha – HMU, Evangelos Markakis – HMU, Konstantina Pityanou – HMU and the rest of the HMU Team, Nikolaus Forgó – UNIVIE, Katarzyna Barud – UNIVIE, Hande Özkayagan-Prändl – UNIVIE, Clara Saillant – UNIVIE, Hillman Velislava – EPL, Maurizio Martignano – UCSC, Christina Farmaki – FORTH, Manolis Spanakis– FORTH, Cristina Regueiro Senderos– TECNALIA, Marisa Escalante – TECNALIA, Rasmusen Sven – FUJITSU, Sofiane Lagraa – FUJITSU, Giulia Rutto – RINA, Ivana Stupar – ENT, Alexandru Nicolin – ISS, Ilias Politis – INQBIT, Kefalas Konstantinos – INQBIT, Vasiliki Maria Sampazioti – INQBIT, Erion Vasilis Pikoulis – ISI / ATHENA, Aris Lalos – ISI / ATHENA, Petros Kapsalas – PASEU, Panagiotis Radoglou-Grammatikis – K3Y, Marios Siganos – K3Y, George Nakas – K3Y, Vasileios Stamatis – K3Y, Maria Zotou – ACCELIGENCE, Nikolaos Papachatzakis – ADR, Akis Kourtis – ZORTENET, Saneea Malik – VTT, Peter Rosengren – CHECKWATT, Sarwar Sayeed – ENU, Nikolaos Pitropakis – ENU**

Dissemination level	PU - Public
Status	Final
Due date	31/07/2023
Document date	31/07/2023
Version number	1.0
	Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

TABLE OF CONTENTS

Index of figures

Index of tables

GLOSSARY OF TERMS

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

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

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

Key Performance Indicators (KPI) - vital metrics used to monitor and evaluate performance in relation to specified goals and objectives offering a scientific approach to performance tracking by providing quantifiable and objective data that can be evaluated and compared over time.

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

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

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

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

Persona - a fictional characterization of a user.

Pilot Campaign - A small or large-scale continuous experimental trial that is used to check, assess, and evaluate the viability of a developed solution and measure its effectiveness.

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

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

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

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

Synthetic data - is artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data. This means that synthetic data and original data should deliver very similar results when undergoing the same statistical analysis. The degree to which synthetic data is an accurate proxy for the original data is a measure of the utility of the method and the model. (https://edps.europa.eu/presspublications/publications/techsonar/synthetic-data_en . The

generation process, also called synthesis, can be performed using different techniques, such as decision trees, or deep learning algorithms. Synthetic data can be classified with respect to the type of the original data: the first type employs real datasets, the second employs knowledge gathered by the analysts instead, and the third type is a combination of these two.

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

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

Usability –the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use. Use case -specification of a set of actions performed by a system, which yields an observable result that is, typically, of value for one or more actors or other stakeholders of the system. [b-IEC 62559-2]

User-centred requirement – a requirement that aims to make the product usable and focused on end user needs and objectives. User Story – a small story created to achieve a particular objective inside a product.

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

LIST OF ABBREVIATIONS AND ACRONYMS

ABB	Architecture Building Block
AI	Artificial Intelligence
AIMaaS	AI Models as a Service
AM	Authentication Manager
ASQ	After-Scenario Questionnaire
ATI	Affinity for Technology Interaction scale
ATR	Accountable Transactions Recorder
CRFs	COVID Case Report Forms
DA	TRUSTEE Dashboard
DL	Data Lake
DMHP	Data Management Handling Plan
DPIA	Data Privacy Impact Assessment
DU	Deep Unrolling
Edtech	Education Technology provider
EU	European Union
ESA	European Space Agency
FAIR	Findability Accessibility Interoperability Reproducibility principles

FC	Facilitating Conditions
FCR-D	Frequency Containment Reserve-Disturbance
FFR	Fast Frequency Response
FL	Federated Learning
GA	Grant Agreement
GDPR	General Data Protection Regulation
HE	Homomorphic Encryption
HEDF	Homomorphic Enabled Data Fusion
HM	Hedonic Motivation
IDS	International Data Spaces
IDSA	International Data Spaces Association
IPFS	InterPlanetary File System
LIDAR	Light Detection And Ranging
KPI	Key Performance Indicator
LMS	Learning Management System
MAPP	Multi Agent Path Planning
ML	Machine Learning
NASA	The National Aeronautics and Space Administration

OneSS	One-Stop-Shop
PoC	Proof of Concept
PV	Price Value
QoE	Quality of Experience
RQ	Research Question
SEND	Student with Educational Needs
SUS	System Usability Scale
SI	Social Influence
SIM	Student Information Management
SSI	Self Sovereign Identity
SSI-HE	SSI Homomorphic Capable Framework
STM	Security and Trust Manager
TAM	Technology Acceptance Model
TAI-SDF	Trustworthy AI Support Design Framework
TLX	Task Load Index scale
UTAUT	Unified Theory of Acceptance and Use of Technology
UX	User Experience
WP	Work Package

XAI	Explainable Artificial Intelligence

EXECUTIVE SUMMARY

This document, D5.1 – "Report on TRUSTEE Pilot Campaign Plan and Evaluation Methodology" is the first deliverable of Work Package (WP) 5 – "TRUSTEE MULTI-DATA SPACES Pilot Campaign" and mainly reports the outcomes of Task 5.1 – "TRUSTEE Pilot Campaign Plan and Evaluation Methodology" that is active since M9 of the project. This deliverable presents an overview of the TRUSTEE Pan-European Pilot Campaign with regard to the objectives, the strategy, the Pilot Campaign Plan, and the Key Performance Indicators (KPIs) of the campaign in general. This deliverable also provides the overall methodology to be followed for the conduction of the entire campaign by developing and providing the TRUSTEE Pan-European Pilot Campaign Definition Framework and the TRUSTEE Pan-European Pilot Campaign Evaluation Framework. The Definition Framework outlines a set of initial steps to be taken for the definition and smooth operation of each Pilot Phase of the Pilot Campaign prior to its execution, while the Evaluation Framework describes an initial set of steps to be followed after the conduction of each Pilot Phase in order to evaluate its results. Furthermore, this deliverable also provides the initial definition of the 1st Pilot Phase of the campaign, namely the "Dry Run Scenario" that starts in the beginning of M13 and runs until M18, using the Definition Framework developed. The evaluation of the 1st Pilot Phase's outcomes will be presented in the next deliverable of WP5, namely D5.2, which is due in M19.

This deliverable also reports preliminary outcomes of Task 5.2 – "Pilot use case 1: ENERGY dataset", Task 5.3 – "Pilot use case 2: HEALTH Dataset", Task 5.4 – "Pilot use case 3: EDUCATION Datasets", Task 5.5 – "Pilot use case 4: AUTOMOTIVE Datasets", Task 5.6 – "Pilot use case 5: SPACE Datasets", and Task 5.7 – "Pilot use case 6: Trusted multi- disciplinary data exchange" with regard to the definition of the 1st Pilot Phase of the campaign. The aforementioned tasks, namely Task 5.2 – Task 5.7, represent the six (6) Pilot Use Cases of the TRUSTEE Project and they are activated at the beginning of M13 while this deliverable is due to be submitted at the end of the same month.

INTRODUCTION

PURPOSE AND SCOPE OF THE DOCUMENT

The TRUSTEE Pan-European Pilot Campaign includes four Pilot Phases and six Pilot Use cases. The four Pilot Phases follow the step-by-step and continuous integration procedure of the various Architecture Building Blocks (ABBs) of the TRUSTEE Platform, which were defined in D2.1 – "Live doc conceptualisation, use cases and system architecture V1" [1], as follows:

- Pilot Phase 1 – "Dry Run Scenario" [M13 – M18]: Reference Architecture Design & Preparation
- Pilot Phase 2 – "Baseline Scenario" [M20 – M25]: Technical Development & Innovation
- Pilot Phase 3 – "Multi-disciplinary Scenario" [M27 – M32]: Prototyping, Integration & Validation
- Pilot Phase 4 – "Cross-sector Scenario" [M34 – M39]: Demonstration, Evaluation & Cost Benefit Analysis.

The six Pilot Use cases of TRUSTEE, namely Energy, Health, Education, Space, Automotive, and Trusted Multi-disciplinary Data Exchange, will run concurrently throughout all Pilot Phases of the entire Pilot Campaign for achieving intercommunication between the different data domains, integration of all ABBs, and, eventually, the delivery of the final prototype of the TRUSTEE Platform by the end of the campaign.

This document defines the steps that need to be taken prior to the run of each of the four Pilot Phases for defining the respective phase to be activated, with regard to:

- Definition of the Use Case Scenarios to be tested during the Pilot Phase
- Identification of Pilot Use case objectives for the respective Pilot Phase
- The initial definition of the KPIs of the respective Pilot Phase, prior to its execution
- Description of the Integration Platform on which the Use Case Scenarios will be tested
- Identification of the potential participants throughout the Pilot Phase
- Definition of the role of each partner of the TRUSTEE consortium for the specific Pilot Phase
- Investigation and/or definition of the Data Acquisition and Exchange procedures that will be adopted during the respective Pilot Phase
- Discussion on the Legal and Socio-ethical Considerations for the Pilot Phase
- Outline of the foreseeable and expected outcomes of the Pilot Phase

By defining and discussing all the aforementioned topics prior to each phase's execution, we aim for the smooth operation of each Pilot Phase and the elimination of unforeseeable risks emerging during pilot testing.

Additionally, this document provides an overview of the entire TRUSTEE Pan-European Pilot Campaign and the initial Pilot Campaign Plan by discussing the objectives of the campaign as well as the plan to be followed, the strategy adopted, the general KPIs set, and the expected outcomes regarding the delivery of the final functional environment.

As part of the activities performed under Task 5.1 – "TRUSTEE Pilot Campaign Plan and Evaluation Methodology", this document also presents an initial version of the TRUSTEE Pan-European Pilot Campaign Evaluation Framework, which will be used for assessing the outputs of pilot testing after the conduction of each Pilot Phase as well as for evaluating the entire campaign towards the end of the project, and which is also subject to updates and enhancements as the project progresses. Several qualitative and qualitative methods and tools are presented in the following sections for data capturing during pilot testing. Data analysis practices have also been considered for extracting knowledge from the data gathered, which will be used for measuring and assessing the realization of the KPIs of each Pilot Phase and of the entire campaign, after its end. User Experience (UX) evaluation methods have also been investigated and are outlined further below in this document as an initial approach to define the TRUSTEE Quality of Exeprience Assessment Framework, which will be further enhanced and adopted after Pilot Phase 2 – "Baseline Scenario" for ensuring that the TRUSTEE solution is user-driven, and the TRUSTEE Platform is tailored around end-user needs and objectives.

RELATION TO OTHER WORK PACKAGES

WP5 is closely related to all other WPs of the TRUSTEE Project. Regarding the activities reported in this document, namely D5.1, input has been received from:

- WP2 regarding the initially defined TRUSTEE Pilot Use Cases, which drive the definition of the Use Case Scenarios to be tested during the various Pilot Phases; the TRUSTEE Personas, which portray the end-users and stakeholders that will be part of the Use Case Scenarios of each Pilot Phase; the User-centred, Legal, Socio-ethical, and Technical Requirements, which will be addressed during pilot testing; the various ABBs of the TRUSTEE Platform's detailed Architecture, by and within which the functionalities to be tested in each Use Case Scenario are developed; preliminary insight for defining the KPIs of the Pilot Phases, with regard to the

ABBs, the TRUSTEE Pilot Use Cases, and the Pilot Campaign as a whole. All the above-mentioned information is reported in D2.1: "Live doc conceptualisation, use cases and system architecture V1".

- WP3 and WP4, as they are the main technical development WPs of the TRUSTEE project, responsible for the implementation and delivery of the functionalities to be tested in the Use Case Scenarios of each Pilot Phase while they also provide KPIs relevant to the various ABBs of the TRUSTEE Platform.

TRUSTEE PAN-EUROPEAN PILOT CAMPAIGN OVERVIEW

An overview of the entire TRUSTEE Pan-European Pilot Campaign is provided in this chapter including the general objectives, strategy, plan, and expected outcomes of the campaign. The Co-deisgn approach adopted under WP2 for the definition of the TRUSTEE Personas and the identification of the User-centred, Legal, Socio-ethical, and Technical Requirements is briefly described with the goal to demonstrate the link between WP2 and WP5. Thorough explanations, complete lists of Personas and of all types of requirements, as well as more information about the Co-design approach can be found in D2.1 [1].

An overview of the six TRUSTEE Pilot Use Cases and their objectives for the entire TRUSTEE Pan-European Pilot Campaign is also provided in this chapter while more details can be found in D2.1 [1]. As part of the overview of the campaign, the four Pilot Phases of the campaign are described in this chapter alongside the time plan and KPIs of the campaign.

PILOT CAMPAIGN OBJECTIVES

The TRUSTEE Pan-European Pilot Campaign aims to assess and deliver the final prototype of the TRUSTEE system through smooth operation testing, by following the step-by-step and continuous integration of the various ABBs of the TRUSTEE Platform. Among the objectives of the entire campaign are the validation of the TRUSTEE ecosystem, with regard to the capabilities and benefits that the TRUSTEE Platform brings to the table for data spaces as a mediator for multi-disciplinary data use, and the impact assessment of the TRUSTEE ecosystem in supporting other European initiatives, such as GAIA-X.

PILOT CAMPAIGN STRATEGY

CO-DESIGN APPROACH FOR PERSONAS, REQUIREMENTS, AND USE CASE SCENARIOS DEVELOPMENT

Co-design is a design approach that seeks to actively incorporate all stakeholders (e.g., partners, stakeholders, end-users) in the design process to guarantee that the result can be efficient and satisfy their needs. Co-design is not a design style, but rather an approach centred on design processes and procedures. It may be used to create environments and solutions that are more receptive and adaptable to the cultural, emotional, spiritual, and practical demands of their end-users [2].

To be more direct and recognize the goals and desires of stakeholders and end-users, the Co-design approach was adopted and utilized for the development of the Personas and User-centred Requirements that led to the identification of the Use Case Scenarios for TRUSTEE, which will be defined and described in this document.

PERSONAS DEVELOPMENT IN WP2

Initial one-on-one interviews were scheduled with the partners leading the TRUSTEE Pilot Use Cases to discuss preliminary technical specifications of the Use Cases and understand the Pilot Leaders' needs from a solution such as TRUSTEE. The results of these interviews prompted the distribution of surveys to the project's Pilot Leaders and Members as thoroughly described in D2.1 [1].

Following the collection and analysis of the responses, a preliminary set of Personas was defined. Following, a Co-design Workshop was scheduled with the goal to, among others, examine, cross-validate, and enhance the developed Personas. The updated TRUSTEE Personas were then integrated and linked to the TRUSTEE Platform's multiple user role categories, yielding the final TRUSTEE Personas, which correspond to the Data Provider User Role, the Consumer User Role, the Model Provider User Role, and the Developer User Role. This procedure is showcased in Figure 1 below.

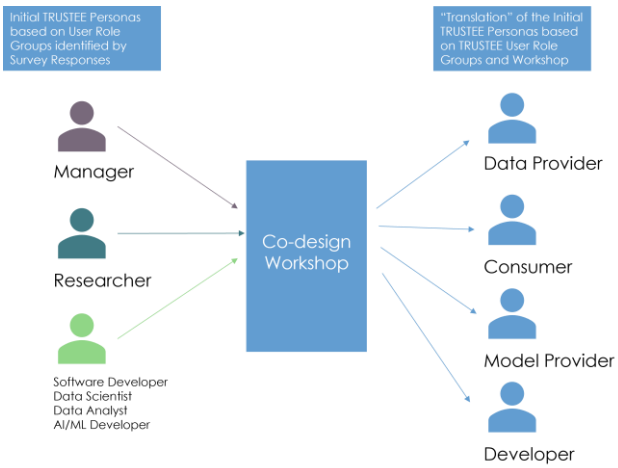

Figure 1: TRUSTEE Developed Personas

The Personas developed under WP2 portray the end-user roles for each TRUSTEE Pilot Use Case. More details and thorough explanations can be found in D2.1 [1].

USER-CENTERED, LEGAL, SOCIO-ETHICAL, AND TECHNICAL REQUIREMENTS ENGINEERING IN WP2

Each Persona includes its goals, frustrations, and ambitions for privacy-preserving data computations as well as secondary data use. This knowledge and insight were used to formulate the User Stories of each Persona [3]. While a Persona showcases the characteristics, goals, and needs of a user, User Stories explore the activities that these users want to accomplish by utilizing the TRUSTEE Platform. Several User stories may stem from a single Persona, and they are then utilized to define the User-centred Requirements. The User-centred Requirements can assist in the definition of the platform's features that are centred on the needs of its end-users, thus, leading to definition of the Technical, functional and non-functional, Requirements. Lastly, once defined, the use of Personas allowed us to further identify and specify the various challenges and requirements to be considered from legal and socio-ethical aspects. As detailed in D2.1 [1], the methodology of mapping and defining the legal and socio-ethical requirements was based on state-of-the-art research and analysis of the currently binding and forthcoming relevant legislature and socio-ethical frameworks relating to the use of Artificial Intelligence (AI) for data manipulation.

The User-centred Requirements that were identified for the TRUSTEE Platform were cross-validated and enhanced during the Co-design Workshop. This process is detailly presented in D2.1 [1] alongside the analysis of its outputs. The User-centred Requirements that were identified portray the objectives and needs of end-users coming from, inter alia, each one of the six Pilot Use Cases of TRUSTEE. Complete lists of User-centred, Legal, Socio-ethical, and Technical Requirements are provided in D2.1 [1].

USE CASE SCENARIOS DEFINITION

A Use Case Scenario describes how a system will be utilized in a real-world setting and defines how the system responds to the steps and interactions that a user will complete when utilizing the system to accomplish their goal. Use Case Scenarios are frequently used in software development and product design to assist in identifying and clarifying the requirements and functionality of a system or product. Furthermore, they may be used to assist stakeholders in understanding how the system will function in practice and identifying potential challenges or points of improvement. The development of Use Case Scenarios can assist designers and developers in assuring that the system can satisfy the demands of its users as those are planned in real-world situations [3].

STRATEGY DEFINITION FOR THE TRUSTEE PILOT USE CASES

PILOT USE CASE 1: ENERGY DATASET

The electricity grid needs a stable frequency of 50 Hz, in Europe, which means that any electricity consumed anywhere in the network also needs to be produced at the same time somewhere in the network.

Various disturbances can occur, for instance, some power plant must suddenly shut down due to malfunction or planned maintenance. Therefore, the grid operator needs to have a backup of resources, often referred to as flexibility resources. These resources will be activated in case of a sudden disruption. Examples of resources are hydropower and batteries. Batteries are mainly involved in providing two flexibility services Fast Frequency Response (FFR) and Frequency Containment Reserve-Disturbance (FCR-D).

The national grid operator in Sweden, Svenska Kraftnät[1], is running an auction every day to secure the next day's need for flexibility in terms of the MWatt effect. Eight (8) authorised bidders are allowed to make bids, and CHECKWATT is one of those authorized bidders. The minimum bid size is 100 kWatt, while the typical daily need is 10-30 MWatt.

CHECKWATT is an aggregator controlling a huge number of batteries, both large (2 MWatt peak effect) and home batteries (5 kWatt peak effect). In order to make a successful bid, first the requested amount of MWatt has to be understood, and second, the need of CHECKWATT's customers' needs to be checked within their database and a prediction is made regarding how weather and other conditions might have an effect on CHECKWATT's available resources. Once a bid is won, the customers' batteries need to be prepared for activation. Finally, when the time for the requested flexibility comes and the frequency goes out of range, all CHECKWATT's resources need to be activated.

[1] https://www.svk.se/en/

The data we will be working with is historical time series of both consumption and production for each customer, spot prices data from NordPool[2], real-time data from batteries (state-of-charge), and weather forecasts.

The following three general objectives are set in the Energy Pilot Use Case of TRUSTEE, which will be further elaborated during the definition of each Pilot Phase of the TRUSTEE Pan-European Pilot Campaign:

- Preparing a bid for the flexibility market
- Monitor and Prepare for Flexibility Activation
- Flexibility Activation
- Investigation on how knowledge extraction can be enhanced by third-party datasets, considering secondary data use

The aforementioned objectives are set with regard to the solution proposed by TRUSTEE, considering data sharing, lightweight computations, and AI/ML models use in a privacy-aware way.

PILOT USE CASE 2: HEALTH DATASET

The TRUSTEE Platform aims to assist researchers from UCSC in identifying the potentiality of utilizing cross-domain data, deploying privacy-enhancing mechanisms such as Homomorphic Encryption (HE), and identifying beneficial data processing by using data resembling health data and data from other domains. The data that will be used within the Health Pilot Use Case will resemble data stemming from COVID Case Report Forms (CRFs) and Contact Tracing Information, as described in D1.6 [4] and D2.1 [1].

The main objectives of the Health TRUSTEE Pilot Use Case for the entire Pilot Campaign are:

- Representing the typical searches/queries that are performed by researchers on the datasets, e.g.:
  - Table searches on numbers and strings,
  - (Partial) Transitive Closure Searches identifying relationships among patients (e.g., positive-case – contact),
  - AI / Machine Learning (ML) based regression analyses to identify relationships among medical/health parameters,
  - Privacy-preserving record linkage to identify the same patients in different datasets.
- Investigate HE-related solutions in terms of performances and size of the data to be processed,
- Investigate the use of Federated Learning (FL) and related solutions in terms of communication, collaboration, and procedural constraints on operations.
- Investigate how knowledge extraction can be enhanced by third-party datasets, considering secondary data use

PILOT USE CASE 3: EDUCATION DATASET

A school environment typically comprises a Student Information Management (SIM) system, which is often hosted in the cloud (by a third-party provider) and more rarely by the school itself (on the school

[2] https://www.nordpoolgroup.com/

premises). The SIM system may be able to connect with a third-party education technology provider (EdTech) who will act as a third-party manager of student data via APIs and provide partial access and sharing of this data of the school on specific students to other third-party providers, such as other EdTech providers, learning and governmental agencies, local school authorities, non-profit organisations providing specific services relating to education. A school would typically also have a Learning Management System (LMS), whose functionalities are to facilitate the learning process by providing a platform for the teaching staff as well as students to manage, deliver content, course work, assessments, communication, and various collaboration and tools for learning. SIMs collect and manage a wide range of student-related data, which can be sensitive and personal – from personal information, enrolment history, "pupil premium" (individuals on reduced lunch fees) or Student with Educational Needs (SEND), attendance records, family information, grades, transcripts, and other data. LMS data also includes some student data, however, its primary focus is on managing academic-related information. It stores and tracks information related to course work, assessments, student progress, and interaction with the learning environment. Some educational institutions may be hosting the LMS on their premises. Typically, however, many would use cloud-based storage offered by the LMS providers. The data from SIMs is of particular interest, in this case, to see because of its sensitivity and for its use and processing by various third parties. The data to be used for the Education Pilot Use Case have been initially described in D1.6 [4] and D2.1 [1] and more information will be provided in upcoming deliverables, namely D1.7, D2.2, and D5.2.

The objectives stemming from providing fictional data from the education domain are three-fold:

- See how sharing data on the TRUSTEE Platform would improve and assure privacy and security in data exchange, use, and manipulation. Dealing with data leakage, security risks, partial identification of data, data transfers, and other similar risks are key aspects that need to be considered in exchanging or sharing partial data so this will be an opportunity to simulate the typical partial data exchange or access, however, in the TRUSTEE case, the objective will be to see how security and privacy mechanisms can be enhanced.
- As a solution, it will be important to see on what part of education data transactions can homomorphic encryption be implemented and what it will protect. For sensitive data, such as children's data, the investigation will be stirred around how HE could be used to enable analysis or sharing of the data, whether it could remove privacy barriers inhibiting data sharing or increasing security to existing services, or if it could reduce some privacy concerns. It is interesting to explore also if it is possible to adhere to all General Data Protection Regulation (GDPR) and if any traceability/compatibility matrix is showing the relationship between the (to be) adopted HE technologies and the (non) satisfied GDPR requirement and how can such techniques be adopted at educational institution level. However, here will also be important to investigate the environmental and other costs HE may incur. Furthermore, it will be an advanced opportunity to investigate how combining education data with other data (such as health or energy or transport data relating to education in some way albeit from other systems and silos) and allow for computation and analysis.
- What controls and at what stage of data use, exchange or computation must be implemented and on which stakeholder (EdTech provider, educational institution, educator - all of them as a condition).

PILOT USE CASE 4: AUTOMOTIVE DATASET

In order to efficiently promote decision-making in an automated or semi-automated manner, Automated driving, traffic management and smart mobility systems demand high volumes of data. Despite the multimodal origin of automotive data, already being gathered today, the explainability of the processing and the usability of this data has been difficult because of technical, legislative, and/or financial aspects. Thus, the challenges stem not only from the need to move high-volume, high-velocity, high-veracity, and diverse data between organizations, but also from industrial competition, complicated administrative procedures, and, most importantly, data protection legislation and limits such as the EU GDPR [5]. Large-scale data gathering and processing at a powerful cloud-based server in standard ML algorithms implies a single point of failure and the possibility of major data breaches. First and foremost, centralized data processing and administration impose limited transparency and provenance on the system, which may result in a lack of confidence from end users as well as trouble complying with the GDPR [6].

In TRUSTEE's Automotive Pilot Use Case, FL [7] has been selected as a strategy for implementing Perception, Prediction and Planning functionalities in a decentralized collaborative learning context, where the algorithm is implemented on numerous local datasets stored at distant agents (i.e., local nodes) vehicles, simulating frameworks and infrastructural sensors, rather than gathering and processing the training data at a centralized data server. In this way, TRUSTEE's framework enables local nodes to train a shared ML model cooperatively while maintaining both the training dataset and computation at internal locations. Only training output (i.e., parameters) are homomorphically encrypted and exchanged at a set frequency, which necessitates the use of a central server to coordinate the training process through a peer-to-peer underlying network architecture (decentralised FL) to aggregate the training results and compute the global model, while periodically updating the local models accommodated on agents through transmitting the homomorphically encrypted global model parameters back to the local agents. This process is illustrated in Figure 2, which exemplifies the sequence of model update and aggregation on edge and cloud levels correspondingly.

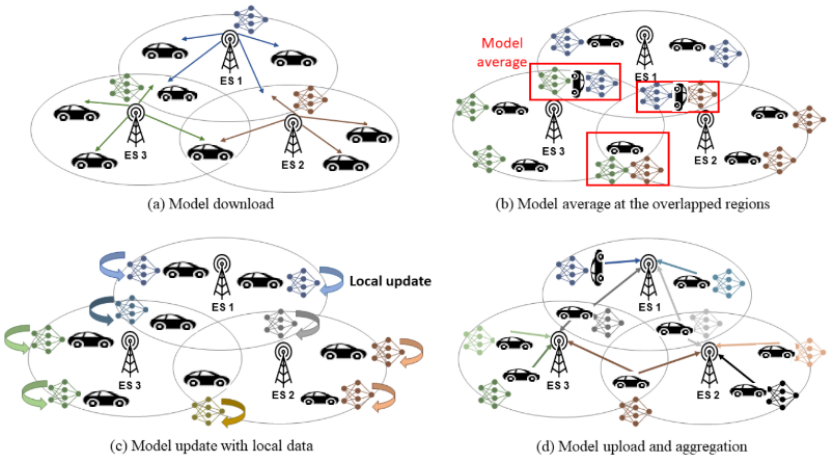

Figure 2: Illustration of the Federated Learning Scheme employed in TRUSTEE's automotive pilot.

Through this approach, the Automotive Pilot Use Case in TRUSTEE establishes an ecosystem of data gathering, environmental modelling and automated navigation by addressing technical, legal, and financial challenges related to:

- Stabilization of ML functions through guaranteeing a high level of data heterogeneity as the data collected by multiple agents highly resemble the scene parameters encountered in the real world.

- Efficient Deployment and management of computational resources, as distributed ML addresses, optimizes the usage of computational resources connected to the network, while suppressing the need for high-volume centralized computational units.
- The exchange of homomorphically encrypted ML model parameters instead of raw data, as would be required in traditional ML schemes, robustifies privacy preservation while enhancing TRUSTEE's platform cyber-resilience.
- Enablement of cooperation between various stakeholders: Industries, Universities, and Governmental institutions without challenging the ownership (sharing) of raw sensing data, which is of high concern for industrial partners, as it is associated with competition.

In addition to the current prevailing concerns in the Automotive Data Spaces, an open data space is also being emerged, that provides access to real-time traffic data as well as automotive data beyond their secure and privacy-preserving transfer and that associates prevalent schemes of data platforms. It is expected that it will be realistic to provide automotive data on the EU level. Relying on a distributed system architecture, proposed and supported by the International Data Spaces Association (IDSA) [8], the Automotive Pilot Use Case in TRUSTEE presents an ecosystem in which pilot leaders, acting as data providers, determine the circumstances ruling the manners and the actors entitled to access and process the data/ metadata. The aforementioned process sets-up data sovereignty and trustworthiness, where data consumers could be confident regarding the origin, integrity and quality of data. Through integrating data from a multitude of modalities and platforms, TRUSTEE's Automotive Pilot Use Case in association with the FL will provide a venue for distributing digital data-driven business models, launching novel schemes for association, exploitation and perception. The Automotive Pilot Use Case will consider all acting parties either data providers, users, developers, or end users.

Within TRUSTEE's platform, the Automotive Pilot Use Case aims to:

- Investigate data sovereignty and security along the chain.
- Investigate the potential of the development of new business models, distribution venues and services, as well as a larger selection of innovative mobility services and applications

The main objectives of this Pilot Use Case concern:

- Incorporation of distributed multimodal fusion for geo-localization in order to increase robustness and resilience.
- Development and testing of four-dimensional situational awareness and quantifying the contribution of the technology on enhancing the level of safety.
- Extension of the potential and limitations of Multi Agent Path Planning (MAPP) towards enhancing the safety metrics and extending the operability on more complex scenarios (supporting more degrees of freedom on the kinematic model of the ego vehicle);
- Exploration of the potential and the limitations of the co-operative awareness engine towards accommodating inputs from multiple agents and multiple sensors; and finally,
- Development and extension of the existing autonomous driving KPIs through introducing application-aware metrics.
- Investigation on how knowledge extraction can be enhanced by third-party datasets, considering secondary data use

The aforementioned objectives are set with regard to the solution proposed by TRUSTEE, considering data sharing, lightweight computations, and AI/ML models use in a privacy-aware way.

PILOT USE CASE 5: SPACE DATASET

The TRUSTEE Space Pilot Use Case was designed to determine the extent to which the TRUSTEE Platform's ability to facilitate the data exchange and processing in the encrypted domain applies to the ISS's scientific mission. Therefore, the technical analysis focuses on the effectiveness of this novel solution for datasets of varying volumes and complexities, as well as for various computing devices. During the experiments, the intention is to track the volume of encrypted data, the complexity of operations in the encrypted domain, and the processing time. The preliminary results of the experiments will therefore be instrumental in future implementations of TRUSTEE's innovative technology, particularly with regard to its potential use in low-power computing solutions (similar to on-board computing) as opposed to high-power computing on the ground segment. In addition to these technical factors, the Space Pilot Use Case will examine a variety of non-technical factors, such as usability, intuitiveness, cross-platform compatibility, etc. As the experiments are intended to utilise publicly available data, with no safety and security constraints and no GDPR constraints, no specific safety and security protocols are planned for the execution of this use case as thoroughly discussed in D1.6 [4] and D2.1 [1].

The data considered for the Space Pilot Use Case pertains to craters and seismic events on the Moon, solar outbursts, and satellite imagery. The datasets on Moon craters and seismic events, as well as those on solar flares, have already been published by well-known international organizations, such as The National Aeronautics and Space Administration (NASA), and are freely accessible, with no GDPR, safety, or security restrictions, so long as they are used for research. The data on seismic events on the Moon is considered classic, in the sense that the first datasets date back to the Apollo 11 mission in 1969, and that this data has been analysed/processed numerous times, typically in conjunction with other related datasets, and is therefore well described in the literature. The vast majority of satellite images originate from the Copernicus programme via the Open Access Hub administered by the European Space Agency (ESA). These images originate from a network of satellites that make a large number of daily observations.

Based on a number of realistic scenarios generated during the early phase of the project, TRUSTEE concentrates on determining how to provide assistance for research and space missions on the ISS. By analysing these, the initial set of objectives for the Space Pilot Use Case includes:

- cross-check the technical and non-technical requirements for the TRUSTEE Platform to ensure that the solution developed is user-driven
- contribute to the development of the framework of the tests that will be carried out in order to verify functioning and user approval, mimicking realistic in-silico experiments and data processing workflows.
- multi-disciplinary cooperation with other fields to identify cross-sector communication to explore cross-correlations with data sets on power shortages and more.
- evaluation of data processing in the encrypted domain by assessing the volume of the encrypted data, the complexity of the operations in the encrypted domain and the computing time

PILOT USE CASE 6: TRUSTED MULTI-DISCIPLINARY DATA EXCHANGE

The objective of the Trusted Multi-disciplinary Data Exchange Pilot Use case of TRUSTEE is to enable secure, standardized, and efficient data sharing across different disciplines and organizations. The pilot will utilize the International Data Spaces (IDS), which is a concept that promotes the secure exchange of data in a decentralized manner, ensuring data sovereignty, privacy, and trust.

Data spaces, particularly those implemented through the concept of IDS, provide a robust framework for secure and efficient cross-border data exchange. IDS fosters data sovereignty, privacy, and trust, making it particularly relevant for cross-discipline and/or international collaborations. By leveraging IDS, organizations can establish a common set of standards and protocols for data exchange, ensuring

interoperability and seamless communication across borders [9]. This allows for the exchange of data between different domains and organizations while maintaining control over data assets. IDS connectors enable secure communication, access control, and authentication, ensuring that only authorized entities can access and exchange data [10].

The standardized data formats supported by IDS facilitate data integration and understanding across borders, removing barriers that may arise from diverse technical specifications. Additionally, IDS emphasizes data governance and compliance, enabling organizations to adhere to relevant regulations and ethical guidelines when sharing data internationally. Through IDS, cross-border data exchange becomes streamlined, transparent, and trustworthy, promoting global collaborations and knowledge sharing across disciplines.

By utilizing the capabilities of the TRUSTEE Platform and leveraging the principles of IDS, the aim of the Trusted Multi-disciplinary Data Exchange pilot is to foster seamless, secure, and trusted collaboration among diverse disciplines and pilots in the TRUSTEE project.

Digital transformation is creating a data ecosystem with data on every aspect of our world, spread across a range of intelligent systems. Consequently, there is a need to bring together data from multiple sources within the data ecosystem. For example, smart cities show how different systems (e.g., energy and transport) within the city can collaborate to maximize the potential to optimize overall city operations. In addition, it is also important to be able to merge data from multiple domains (cross-domain data exchange).

Multi-disciplinary data exchange across domains offers immense potential for innovation, efficiency, sustainability, and safety. By breaking down silos and promoting collaboration, organizations can leverage diverse data sources and knowledge to tackle complex problems, drive advancements, and improve outcomes in sectors such as health, automotive, space, and energy. In Table 1, some of the key benefits of the multi-disciplinary data exchange are listed.

Table 1: Benefits of the multi-disciplinary data exchange

Accelerated innovation	Multi-disciplinary data exchange encourages the sharing of knowledge, expertise, and insights across domains. This cross-pollination of ideas fosters innovation by applying techniques, methodologies, and best practices from one domain to another. For example, advancements in space technology can inspire new approaches in the automotive or energy sectors, leading to accelerated innovation and technological advancements.
Enhanced problem-solving	Combining data from different domains enables a more comprehensive understanding of complex problems. By integrating data from the health, automotive, space, and energy sectors, interdisciplinary collaborations can uncover interdependencies, patterns, and correlations that might not be apparent when examining data within individual domains. This holistic perspective facilitates more effective problem-solving and decision-making.
Improved efficiency and resource optimization	Multi-disciplinary data exchange helps optimize resource utilization. For instance, sharing data and insights between the automotive and energy sectors can facilitate the development of electric vehicles, enabling energy-efficient transportation solutions. Similarly, leveraging health data in the automotive domain can aid in designing safer and more ergonomic vehicles. By exchanging data across domains, organizations can minimize redundancies, streamline processes, and optimize resource allocation.

Enhanced safety and sustainability	Multi-disciplinary data exchange can contribute to improved safety and sustainability practices. For instance, sharing health data with the automotive industry can help identify potential risks or health-related issues associated with vehicle design or usage. Data exchange between the energy and space sectors can promote sustainable energy solutions and space exploration technologies that minimize environmental impact. By collaborating and exchanging data, industries can work together to address safety concerns and promote sustainable practices
Data-driven insights and predictive analytics	Integrating data from different domains can unlock valuable insights and enable predictive analytics. For example, combining health data with automotive and energy data can help predict health-related risks associated with pollution levels, driving behaviours, or energy consumption patterns. These insights can inform policy-making, resource planning, and preventive measures to improve public health, reduce environmental impact, and enhance overall well-being
Cost reduction and optimization	Multi-disciplinary data exchange can lead to cost reduction and optimization across domains. By sharing data, research findings, and experiences, organizations can avoid duplicative efforts and leverage existing resources. For instance, automotive companies can benefit from space industry advancements in materials science, reducing research and development costs. Similarly, energy sector optimization strategies can be applied to healthcare facilities to enhance energy efficiency and reduce operational costs.
Collaborative research and development	Multi-disciplinary data exchange encourages collaboration and joint research and development efforts. For instance, health data can contribute to the development of personalized medicine in collaboration with the pharmaceutical industry. Space technology advancements can aid in the development of satellite-based solutions for energy monitoring and optimization. By combining resources and expertise from different domains, collaborative research and development can lead to breakthrough innovations and solutions that address complex challenges.

The main objectives of this Pilot Use Case include:

- Enable secure, standardized, and efficient data sharing across different disciplines and organizations
- Investigate how knowledge extraction can be enhanced by fostering multi-disciplinary data exchange, considering secondary data use

The aforementioned objectives are set with regard to the solution proposed by TRUSTEE, considering data sharing, lightweight computations, and AI/ML models use in a privacy-aware way.

PILOT CAMPAIGN PLAN

STUDY DESIGN

By adopting a step-by-step and continuous integration of the various ABBs of the TRUSTEE Platform [11] and by splitting the development into 6-month cycles, namely Pilot Phases 1-4, TRUSTEE executes an incremental co-creation development process. Each phase corresponds to other WPs and a set of tasks within the project's overall scope: WP3 and WP4 are the main technical development WPs of the TRUSTEE project, which feeds WP5 with the functionalities to be demonstrated and evaluated during pilot testing. The development and validation of technological components are motivated by WP2's user needs, goals, and requirements, and are carried out in full compliance with EU and national legal, ethical, and fundamental rights frameworks as part of WP1 and WP6's horizontal activities spanning the project's entire lifecycle.

The incremental co-creation development process is further realized by WP5 identifying through pilot testing potential enhancements which are then provided back to WP2, WP3, and WP4, ensuring that the TRUSTEE Platform is tailored around end-user needs and objectives. For each Pilot Phase there will be a clear time plan followed, an integration platform used to ensure the smooth operation of the Pilot Phase, while participants will be identified, as well as clear partner roles and responsibilities will be assigned for ensuring that goals are achieved.

PILOT PHASES

Each phase is delivering results and outcomes that will offer vital input to the phase following, while the cohesive work and close collaboration of all the corresponding partners ensure the continuous integration foreseen in the project. The description, as well as the scope and objectives of each Phase, will be elaborated in the following subsections.

PHASE 1 – DRY RUN SCENARIO: REFERENCE ARCHITECTURE DESIGN & PREPARATION

The design and preparation phase marks the beginning of the project and covers the elaboration of the reference architecture, as well as the large pilots' design for the project and the selection of the technologies to be used. More specifically, the main activities upon which Pilot Phase 1 is based are described in D2.1 [1] and include the following:

- **Identification of the state-of-the-art** relevant to the objectives of the TRUSTEE project. This is accomplished through the investigation of the know-how of the project partners, as well as the exploration of the best practices from existing data service solutions available and the relevant current research bibliography, as reported in D2.1 [1].
- **Definition of User-centred, Legal, Socio-ethical, and Technical requirements**, based on real-world user case scenarios, as previously described. The User-centred requirements are in turn translated into Technical requirements, focusing on the aspects of functionality, usability, reliability, security, performance efficiency, compatibility, maintainability and portability. Additionally, this step considered a detailed analysis of the social, ethical and privacy implications of data acquisition and distribution technologies, from which an initial set of relevant **Legal** and **Socio-ethical requirements** emerged. However, it is imperative that the legal requirements are constantly updated through the lifecycle of the project, as new regulations constantly emerge, such as the upcoming European AI Act, which will ensure a human-centric and ethical development of Artificial Intelligence in Europe. Further details can be found in D2.1 [1].
- **Definition of the reference system architecture,** which comprises all the core system subcomponents and their interconnections that are thoroughly described in D2.1 [1].

Based on the above, the main activity to be carried out during the 1st Pilot Phase considers the **Fast Proof of Concept (PoC) implementation** which will be demonstrated through mock-up prototypes of the various subsystems, following the incremental deployment strategy.

The definition of the technical requirements, the reference architecture, and the most appropriate technologies to be used are carried out within WP2 as they are necessary for the development of the individual subcomponents of the TRUSTEE solution under WP3 and WP4. The 1st Pilot Phase foresees the release of a mock-up architecture and during this phase, the PoC and the mock-up prototypes will begin to be made available. The integration process of the various ABBs of the TRUSTEE Platform will

be carried out in the following Phases of the campaign; however, investigation towards integration will be initiated earlier, as soon as the Pilot Campaign starts, with the goal of smooth operation in subsequent phases.

PHASE 2 – BASELINE SCENARIO: TECHNICAL DEVELOPMENT & INNOVATION

This phase takes off from the User-centred requirements specifications and is composed of a range of integrated, multidisciplinary research and technology tasks. The 2nd Pilot Phase will also translate the Mock-up prototype and PoC from the 1st Pilot Phase into an Initial Working prototype early in the development phase concluding to the final Working prototype that will be provided to the next phase. The step-by-step and continuous integration procedure will accommodate continuous integration cycles stemming from the mock-up and PoC and concluding with the working prototype with as many iterations as needed during Pilot Phase 2. The main procedure to be followed is described below:

- **Identifying technology specifications** from data stakeholders' requirements specifications is the first important task. Also, basic requirements for the data market, etc. brokering of products and tools that are necessary to develop the TRUSTEE ecosystem are foreseen to be identified.
- **Investigating current offerings** encompasses research of existing technologies available from consortium partners, commercial off-the-shelf vendors and other EU, international and/or (inter-) national research projects.
- **Analysing gaps and specifying research needs** involves a comparison of needed and available technologies leading to a set of specifications for new research and development demands as well as a catalogue of the Best Available technologies concerning the innovative technologies addressed in the TRUSTEE project.
- **Defining a prototype architectural platform** involves performing a conceptual design of the TRUSTEE Platform considering the core modules and the data services along with their high-level functional, technical and interoperability specifications. The architecture definition process will address the whole information security situational awareness procedure.
- **Designing a business environment** that can provide the framework for the exploitation of the TRUSTEE prototype architecture foreground and its exploitable products. Extensive market analysis, technology assessment, and business modelling are involved for successful penetration in the emerging area of data market empowering green economy and data development.
- **Design and Development of the architectural elements** that will comprise the proposed awareness ecosystem, fully tailored to the needs of the data economy and user requirements set in the previous phase.
- **Data collection** covers the actual execution of the data optimization algorithms benchmarking tests and the actual collection of data.

The activities included in the 2nd Pilot Phase are performed in close collaboration with the teams in WP3 and WP4, developing the ABBs of the TRUSTEE Platform, and mirror the technical progress in these WPs.

PHASE 3 – MULTI-DISCIPLINARY SCENARIO: PROTOTYPING, INTEGRATION & VALIDATION

With the successful completion of all tasks in the previous phases, the project will have reached the stage, where the realization of the TRUSTEE system in a real-world environment is possible, via the following activities:

- **Integrated architecture development** involves the design, manufacture and assembly of various toolkits and data services. The reference architecture along with the data service and events will be used both for experimental verification in the research phase and for integration into the TRUSTEE reference architecture. While the step-by-step and continuous integration cycles that will be used in the previous phases will have already run the basic testing, the continuous integration cycles will need to acquire the requirements stemming from the integration of the architecture. Integration tests will be part of this phase to identify potential leaks and bugs in the prototype system before its installation and evaluation in realistic conditions.
- **Prototype integration and testing** involves the integration of the project's enabler into the real-world application environments that the pilots will provide to the system, to validate the interoperability among diverse and various data services and the reference architecture interoperability with actual data grid systems.

The integrated version of the TRUSTEE system is foreseen to be implemented under WP3 and WP4 and validated within T5.1 of WP5.

PHASE 4 – CROSS-SECTOR SCENARIO: DEMONSTRATION, EVALUATION AND COST-BENEFIT ANALYSIS

Due to the step-by-step and continuous integration approach adopted by TRUSTEE, this phase will be run almost in parallel with design, development, and integration activities, and will focus on fine-tuning and validation of the whole framework as well as on the assessment of the demonstration phase of the project. Thus, this phase of the project will be concerned with the iterative deployment of the envisioned data enabler architecture to the business scenarios of TRUSTEE (WP6) as well as with the overall project evaluation (lessons learned) and, in the sequel, the preparation of activities regarding the sustainability of the project achievements. Overall actions will include activities such as:

- **TRUSTEE Platform Acceptance** involves the execution of tests, the recording of findings and the addressing of identified shortcomings. Furthermore, Lab integration tests and Simulations will be conducted to identify potential leaks and bugs in the prototype system before its deployment and evaluation in realistic conditions.
- **Validation** involves the final validation of the whole reference architecture against the data stakeholder's requirement specifications and the developed business and exploitations plans.
- **Evaluation** of the entire project and its foreground along with tangible achievements compared to the initial project objectives, with adequate focus on technical evaluation (i.e., KPIs), data stakeholders' acceptance and impact assessment.

TIME PLAN

The time plan of the Pan-European Pilot Campaign of TRUSTEE, as can be seen in Figure 3 below, foresees five stages of pilot testing during each Pilot Phase, for the smooth operation of the entire campaign and each respective Pilot Phase, and the continuous integration of the various ABBs of the TRUSTEE Platform, which will result in the delivery of the final prototype. The stages that will comprise each Pilot Phase are the following:

- Preparation – Use Case Scenario Definition stage, in which the respective Pilot Phase will be defined including the identification of the Use Case Scenarios to be tested; the results of this stage will be reported in respective deliverables, namely D5.1, D5.2, D5.3, D5.4.

- Pre-Demo - Use Case Scenario pre-testing stage, in which a demonstration of the functionalities to be tested in the defined Use Case Scenarios of each Pilot Phase will be demonstrated for eliminating unforeseeable risks during the main pilot testing stages.
- Demonstration – Use Case Scenario testing stage, in which all the Use Case Scenarios defined for the respective Pilot Phase will be tested.
- Post Demo – Phase Evaluation stage, in which the results and outcomes of the respective Pilot Phase will be assessed and KPIs will be measured and reported in the respective upcoming deliverables, namely D5.2, D5.3, D5.4, D5.5.

As can be seen in the following figure, the Preparation – Use Case Scenario Definition stage is foreseen to run just before the initiation of each Pilot Phase.

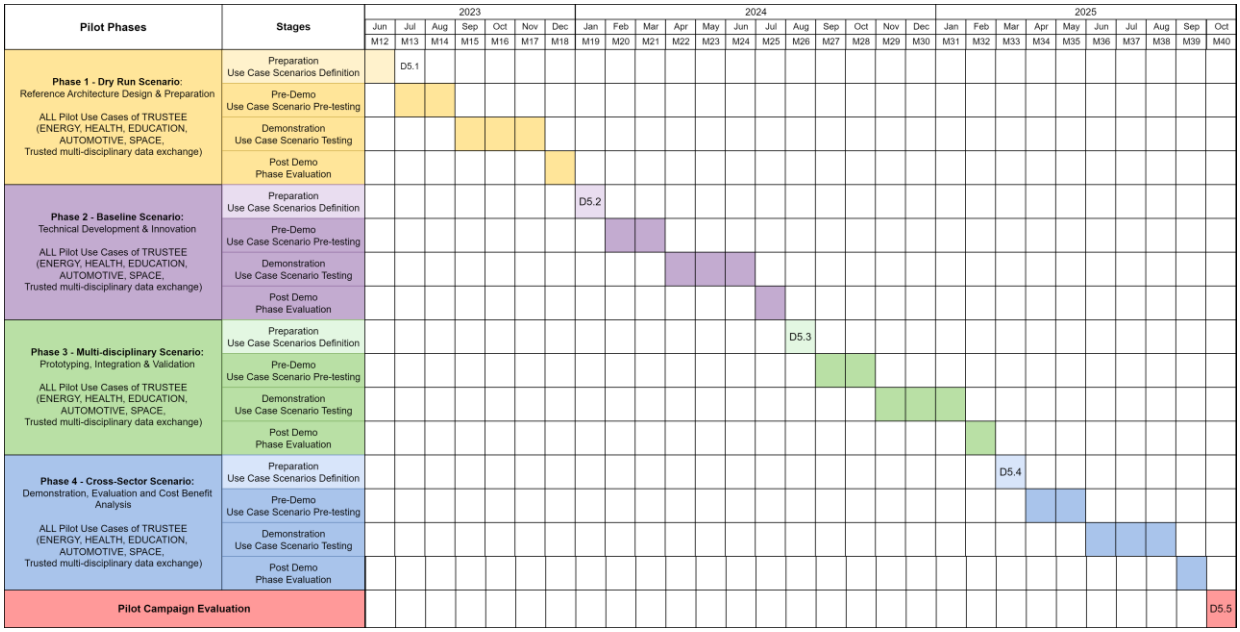

Figure 3: TRUSTEE Pan-European Pilot Campaign Time Plan

The initial approach adopted for reporting the definition of each Pilot Phase and their respective results and evaluation throughout the entire TRUSTEE Pan-European Pilot Campaign can be observed in above Figure 4 and is reflected as follows in the deliverables to be submitted under WP5:

- D5.1 – "Report on TRUSTEE Pilot Campaign Plan and Evaluation Methodology" [M13]: Development and presentation of the Definition Framework and the Evaluation Framework of the TRUSTEE Pan-European Pilot Campaign & Definition of the 1st Pilot Phase by using the Definition Framework defined.
- D5.2 – "Live document on TRUSTEE Pilots V1" [M19]: Results reporting and Evaluation of the 1st Pilot Phase by using the Evaluation Framework defined & Definition of the 2nd Pilot Phase
- D5.3 – "Live document on TRUSTEE Pilots V2" [M26]: Results reporting and Evaluation of the 2nd Pilot Phase & Definition of the 3rd Pilot Phase
- D5.4 – "Live document on TRUSTEE Pilots V3" [M33]: Results reporting and Evaluation of the 3rd Pilot Phase & Definition of the 4th Pilot Phase
- D5.5 – "Live document on TRUSTEE Pilots V4" [M40]: Results reporting and Evaluation of the 4th Pilot Phase & Results and Evaluation of the entire TRUSTEE Pan-European Pilot Campaign

The above-outlined approach is an initial methodology foreseen to be used during the campaign; however, it may be revised and improved depending on additional demands that may arise during the project and the campaign itself.

KEY PERFORMANCE INDICATORS (KPIS)

KPIs are vital metrics used to monitor and evaluate performance concerning specified goals and objectives. KPIs offer a scientific approach to performance tracking by providing quantifiable and objective data that can be evaluated and compared over time and they offer a powerful tool that assists in focusing on what is truly important, ensuring that resources and efforts are consistent with the strategic objectives. Furthermore, KPIs can establish accountability and transparency, while by observing and reporting on KPIs, commitment to attaining goals may be demonstrated, the efficacy of solutions and policies can be evaluated, and progress can be overseen, regulated, and reported.

Moreover, KPIs are crucial for fostering continual development and innovation. Monitoring KPIs allows the detection of trends, opportunities, and issues that may necessitate novel approaches or strategies. Consequently, this may result in the adoption of additional processes that improve performance and competitiveness.

For the Pilot Evaluation process of the TRUSTEE Platform, three different sets of KPIs will be considered. The first set relates to the General KPIs which will be established in all TRUSTEE Use Cases; the second relates to the KPIs Per Pilot Use Case; the third set relates to the KPIs Per Pilot Phase. The first two sets, namely the General KPIs set and the Per Pilot Use Case KPIs set were initially described in the Grant Agreement of the TRUSTEE Project, while the latter set will be defined along the course of the project and before each Pilot Phase, respectively. Below, the first two sets of KPIs are further discussed, namely the set of General KPIs and the set of Pilot Use Case KPIs, and the respective target values are presented. These sets of KPIs will be evaluated at the end of the TRUSTEE Pilot Campaign.

GENERAL

The KPIs that refer to the TRUSTEE project as a whole and will be assessed at the end of the project, stemming from the Grant Agreement (GA), are described in Table 2 below.

Table 2: KPIs of the TRUSTEE Project

KPI #	KPI	Value
G1	Reduce Cost Improvement from all the above; production cost improvement for the end user (industry)	5%
G2	Reduction of Environmental Hazards, by effective data management	10%
G3	Responsible/trustworthy AI	User satisfaction
G4	Integrating scientific knowledge and accurate cross-sector data	2 Data sources concurrent
G5	Optimising/minimising/de-centralising processing, transfer and storage of data and avoiding unnecessary data manipulations	Stakeholder involved

G6	Added value data products, due to qualitative improvement and quantitative production increase	2 Products
G7	Percentage of Improvement of Environmental Impact through reduction and better use of data sources	5%
G8	Win-Win collaboration with Data producers, by providing better services and building a robust partnership	Stakeholder involved
G9	Proactive control on production risks, by enabling easy identification and effective address.	2 Use Cases
G10	facilitate sharing and manipulation of data in compliance with prevailing and emerging legislation (e.g., GDPR)	Stakeholder involved

PER PILOT USE CASE

The KPIs per Pilot Use Case refer to each Pilot Use Case as a whole and will be assessed at the end of the project, after the completion of the entire TRUSTEE Pan-European Pilot Campaign. Similar to the General KPIs, this set is stemming from the Grant Agreement. Following, the KPIs per Pilot Use Case are provided in Table 3, Table 4, Table 5, Table 6, Table 7, and Table 8 regarding the Energy Pilot Use Case, the Health Pilot Use Case, the Education Pilot Use Case, the Automotive Pilot Use Case, the Space Pilot Use Case, and the Trusted Multi-disciplinary Data Exchange Pilot Use Case, respectively.

Table 3: KPIs of the Energy Pilot Use Case

Pilot Use Case 1: ENERGY Dataset		
KPI #	**KPI**	**Value**
UC1.1	New customer services implemented and offered	2
UC1.2	Ensure user-friendly, safe, trustworthy, compliant, fair, transparent, accountable	Rate 0-10
UC1.3	Achieve People satisfaction in the following processes (user-friendly, safe, trustworthy, compliant, fair, transparent, accountable)	70%
UC1.4	Ensure interoperability and reasonable re-use of common reference models	At least 1

Table 4: KPIs of the Health Pilot Use Case

Pilot Use Case 2: HEALTH Dataset		
KPI #	**KPI**	**Value**
UC2.1	Identify cases of data multi-disciplinarity	2
UC2.2	Ensure user-friendly, safe, trustworthy, compliant, fair, transparent, accountable	Rate 0-10

UC2.3	Achieve People satisfaction in the following processes (user-friendly, safe, trustworthy, compliant, fair, transparent, accountable)	70%

Table 5: KPIs of the Educational Pilot Use Case

Pilot Use Case 3: EDUCATIONAL Dataset		
KPI #	**KPI**	**Value**
UC3.1	Ensure user-friendly, safe, trustworthy, compliant, fair, transparent, accountable	Rate 0-10
UC3.2	Achieve People satisfaction in the following processes (user-friendly, safe, trustworthy, compliant, fair, transparent, accountable)	70%
UC3.3	Ensure interoperability and reasonable re-use of common reference models	At least 1
UC3.4	Facilitate use and validation of the TRUSTEE Platform	Up to 500 People
UC3.5	Ensure data subjects/rightsholders and other stakeholders	>90%

Table 6: KPIs of the Automotive Pilot Use Case

Pilot Use Case 4: AUTOMOTIVE Dataset		
KPI #	**KPI**	**Value**
UC4.1	Quality of the 3D scene understanding provided by AD functions	On/Off
UC4.2	Quality of the Emergency alert derived by the AD function	On/Off
UC4.3	Driver's Awareness in terms of scene understanding	On/Off
UC4.4	Driver's ability to control the vehicle in dynamic scene changing (traffic agents crossing the ego-vehicles path)	On/Off
UC4.5	Achieve People satisfaction in the following processes (user-friendly, safe, trustworthy, compliant, fair, transparent, accountable)	70%

UC4.6	Ensure interoperability and reasonable re-use of common reference models	At least 1

Table 7: KPIs of the Space Pilot Use Case

Pilot Use Case 5: SPACE Dataset		
KPI #	**KPI**	**Value**
UC5.1	Researchers and computer scientists at ISS who will take part in the TRUSTEE Pan-European Pilot Campaign	Up to 20
UC5.2	Multi-disciplinary cooperation with other fields to identify cross-sector communication	At least 1
UC5.3	Ensure interoperability and reasonable re-use of common reference models	At least 1
UC5.4	Ensure data subjects/rightsholders and other stakeholders	More than 90%

Table 8: KPIs of the Trusted Multi-Disciplinary Pilot Use Case

Pilot Use Case 6: Trusted Multi-disciplinary Data Exchange		
KPI #	**KPI**	**Value**
UC6.1	Validate trust, reputation, and cooperation mechanisms data providers in the TRUSTEE ecosystem, including at confidence-focused parameters included in SSI e.g., successful delivery of computation, total time of computation and reliability of resource (measured as a percentage of uptime)	3 data providers 3 confidence focused parameters
UC6.2	Achieve People satisfaction in the following processes (user-friendly, safe, trustworthy, compliant, fair, transparent, accountable)	70%
UC6.3	Ensure interoperability and reasonable re-use of common reference models	At least 1
UC6.4	Facilitate the use and validation of the TRUSTEE Platform	Up to 500 People
UC6.5	Ensure data subjects/rightsholders and other stakeholders	More than 90%

PILOT CAMPAIGN EXPECTED OUTCOMES

The expected outcomes of the TRUSTEE Pilot Campaign are based on the Campaign objectives, as well as the defined KPIs. In this section, the focus is on the general expected outcomes of the entire Pilot Campaign, which will be evaluated when all phases of the Pilot Campaign will be completed. In the section "Expected Outcomes of Pilot Phase 1", provided further below in the "Definition of Pilot Phase 1: Dry Run Scenario" chapter, we present expected outcomes with a focus on the 1st Pilot Phase of the TRUSTEE Pan-European Pilot Campaign.

The outcomes were identified and grouped in the following areas: adoption and usage, data economy growth, scalability and integration, security and data privacy, and training and knowledge transfer. We discuss each of the outcome areas and expected outcomes of the Pilot Campaign in the sections below.

ADOPTION AND USAGE

The key aspect of a system's success in the realization of its goals and objectives is the adoption and usage of the system. Ensuring a good user adoption rate and encouraging users to use the system to its full potential will lead to better product quality based on user feedback and continuous improvement. This requires that the platform and its components are adopted to real user requirements and realistic use case scenarios in early development phases.

To achieve and maintain the favourable adoption of the TRUSTEE Platform by users, new customer services and added-value data products will be offered by pilots, ensuring the utilization of TRUSTEE's features. TRUSTEE outcomes such as reducing the environmental impact by better and more efficient use of data, and responsible and trustworthy AI will contribute to user satisfaction. TRUSTEE will increase user adoption by improving the trustworthiness and transparency of the AI models, via the employment of Deep Unrolling (DU) techniques for designing and training interpretable, and computationally and data-efficient AI models. In addition to the available data sources and models, the tools and handbooks, tutorials, and guidelines offered by TRUSTEE will improve the efficiency of the developer stakeholders, leading to production cost improvement, and hence also facilitating the adoption rate among end users.

DATA ECONOMY GROWTH

TRUSTEE will ensure the interoperability and reusability of data from various data sources and domains, as well as European data spaces while relying on the FAIR principles ensuring the data is Findable, Accessible, Interoperable, and Reusable. This will facilitate sharing of data in a trustworthy, privacy-preserving, and reliable way, in compliance with relevant legislation, which will lead to collaboration with data producers and build strong partnerships among TRUSTEE stakeholders. In addition to promoting strong collaboration between data producers and industries participating in the data economy, TRUSTEE Pilot Campaign will demonstrate multidisciplinary, cross-sector communication and data access which will enable the development of new services and products as well as wider and more effective usage of data. Also, the platform will allow developers and data producers to implement and offer new value-added customer services without compromising security and privacy. Additionally, the TRUSTEE Platform will, due to the provided support for developers and facilitating access to data and AI/ML models, lead to more efficient and cost-effective development. The synergy of stated effects will have a positive impact on the development of the European data economy.

SCALABILITY AND INTEGRATION

During the Pilot Campaign, integration and scalability of the TRUSTEE Platform will be assessed and validated. In order to ensure that TRUSTEE integration will provide seamless operation and data exchange, tests will be developed to assess and validate both functionalities of each ABB involved in the pilot and use case scenarios, as well as the integrated system. Scenarios for a stress test of infrastructure resilience, scalability, and user acceptance tests will be conducted. The platform and its components will be validated with realistic amounts of data. Integration of all components will result in a functional environment that will be able to deliver all required functionalities and services and the final prototype of the TRUSTEE system will be delivered. The multitude of data acquisition modes in association with the synergy of geographically distributed agents is crucial for integrating data of high heterogeneity, resembling real-world conditions, which is extremely important for some of the pilots, e.g., Automotive pilot in order the safety indexes not to be disturbed while TRUSTEE's platform to also be scalable.

SECURITY AND DATA PRIVACY

One of the key TRUSTEE's objectives is to enable a trustworthy and privacy-preserving exchange of data. Through TRUSTEE's privacy-by-design architecture and by applying techniques such as HE, FL, and compliance with legislations and regulations such as the GDPR, TRUSTEE Pilot Campaign will demonstrate secure and privacy-preserving sharing and manipulation of data. Additional security will be facilitated by optimizing and decentralizing the processing, transfer, and storage of data.

TRAINING AND KNOWLEDGE TRANSFER

An important part of TRUSTEE's adoption among users is providing training and knowledge transfer to involved stakeholders. Informative and well-organized knowledge transfer including comprehensive and high-quality documentation and training sessions will be performed so that end-users can get introduced to the platform's features and learn how to effectively use the TRUSTEE Platform. After providing such education and training sessions, the maximal exploitation of the platform can be realized. The system will be validated through interaction with the users and feedback will be gathered during training.

TRUSTEE PAN-EUROPEAN PILOT CAMPAIGN DEFINITION FRAMEWORK

In order to ensure smooth operation of the Pilot Phases, the Definition Framework includes the initial set of guidelines to be used prior to the initiation of each Pilot Phase for defining the respective Phase. This set of guidelines constitutes the TRUSTEE Pan-European Pilot Campaign Definition Framework and are further explained in the sections that follow.

PILOT PHASE SUMMARY

A summary will be provided for each Pilot Phase offering the outline of its definition according to the guidelines of the Definition Framework.

PILOT PHASE OBJECTIVES

A set of Pilot Phase specific objectives will be defined prior to the execution of each Pilot Phase to ensure that the goal of the Pilot Phase is clear. These objectives will pave the way towards assessing the general

outcome of the respective Pilot Phase after its completion alongside the assessment of Pilot Phase specific KPIs.

PILOT PHASE USE CASE SCENARIOS

Regarding the platform of TRUSTEE, the Use Case Scenarios of each Pilot Phase will be defined by extracting knowledge from the developed Personas, User Stories, and User-centred, Legal, Socio-ethical, and Technical Requirements, as defined in D2.1 [1]. All types of requirements that will be associated or addressed in a certain extent through the defined Use Case Scenarios by taking into consideration the maturity of the various ABBs during the respective Pilot Phase, which is strongly linked with the step-by-step and continuous integration approach adopted in TRUSTEE. The following template will be used for the synthetic and condensed definition of the Use Case Scenarios to be tested during the four Pilot Phases of the TRUSTEE Pan-European Pilot Campaign, as presented in Table 9.

Table 9: Use Case Scenario Template

Use Case Scenario 00X	
Scenario ID	*Format: UCS-00X* *A unique ID for the Use Case Scenario*
Title	*A title for the scenario*
Scenario Description	*A brief description of the Scenario and the functionalities to be tested*
Digital Solutions/Functionalities to be tested	*Format: ABBName:FunctionalityName/ID* *e.g., AM:Authentication/FNC-00X*
Technical partners involved	*Partners' Short Names/Abbreviations* *e.g., HMU*
ABBs and/or inner modules of ABBs that are related	*Format: ABB name and/or ABBName:InnerModuleName* *e.g., Dashboard, Dashboard:RPA*
User-centered Requirements	*The User-centered Requirements that are involved in/addressed by the scenario.* *e.g., USR-00X*
Technical Requirements	*The Technical Requirements that are involved in/addressed by the scenario.* *e.g., Req-AIMaaS-FUNC-1*
Legal Requirements	*The Legal Requirements that are involved in/addressed by the scenario.* *e.g., LEG-TSD-00X, LEG-USR-00X*
Ethical Requirements	*The Socio-ethical Requirements that are involved in/addressed by the scenario.* *e.g., SOC-ETH-00X*
Piloting Summary	*In which Pilot Phase(s) the scenario will be tested and within which Pilot Use Case(s)*
Main Persona(s) involved in the scenario	*Format: PersonaName:FunctionalityName/ID* *e.g., DataProvider:Authentication/FNC-00X*
Comments / Open issues	*Any comments related to the scenario, potential outcomes that are anticipated, or open issues that are foreseen to might emerge during the scenario.* *e.g., if the scenario is to be tested in more than one phase, then indicate the open issues that still need to be addressed from phase to phase*

PILOT PHASE STRATEGY

A clear strategy will be defined for each Pilot Phase including the Use Case Scenarios to be tested within each one of the six TRUSTEE Pilot Use Cases alongside the objectives of each Pilot Use Case for the respective Pilot Phase and the methodology to be followed. The Use case scenarios to be tested in Phase 1 will be outlined in this deliverable in the "Definition of Pilot Phase 1: Dry Run Scenario" chapter, while the Use Case Scenarios of the Phases 2, 3, and 4 will be defined in D5.2, D5.3, and D5.4, respectively. Table 10 below provides the template to be used for outlining the objectives of each Pilot Use Case for the respective Pilot Phase.

Table 10: Pilot Use Case Objectives Definition Template

Pilot Phase X – Title			
Objective #	**Objective**	**Pilot Domain**	**Use Case Scenario**
Objective ID (Format: PPX.UC.ObjX)	*Description of the Objective*	*Pilot Domain that is related to the objective*	*Use Case Scenario with which the Objective is related*

PILOT PHASE KPIS

As mentioned in the "TRUSTEE Pan-European Pilot Campaign Overview" chapter, three sets of KPIs are defined for the TRUSTEE Pan-European Pilot Campaign, namely: the General KPIs relevant to the campaign as a whole, the KPIs that refer to each Pilot Use case of the TRUSTEE project, and the KPIs that will be defined prior to each Pilot Phase. The latter set of KPIs will be defined by the partners leading the functionalities to be demonstrated in each Use Case scenarios in collaboration with the partners leading the Pilot Use Cases of TRUSTEE and will be evaluated after the conduction of the respective Pilot Phase. Pilot Phase KPIs will include technical-orientated KPIs for the functionalities developed within the ABBs of the TRUSTEE Platform as well as Pilot Use Case KPIs specific for each TRUSTEE Pilot Use Case in the context of each Pilot Phase as the project progresses.

The KPIs of 1st Pilot Phase of the TRUSTEE Pilot Campaign are defined in this deliverable (D5.1) further below in the "Definition of Pilot Phase 1: Dry Run Scenario" chapter of this document. Regarding the rest of the Phases, namely Pilot Phases 2, 3, and 4, the corresponding KPIs will be defined in deliverables D5.2, D5.3, and D5.4, respectively. The template presented in Table 11, below, will be used for the definition of Pilot Phase KPIs.

Table 11: KPIs Definition Template

Pilot Phase X – Title					
KPI #	**KPI**	**ABB**	**Pilot Domain**	**Impact**	**Value**
KPI ID (Format: PPX.X)	*Description of the KPI*	*ABB for which the KPI is defined*	*Pilot Domain in which the KPI will be evaluated*	*Impact of the KPI*	*Target Value of the KPI*

PILOT PHASE INTEGRATION PLATFORM

Pilot tests will be conducted by the project's collaborating partners while the TRUSTEE solution is foreseen to be demonstrated and verified in real operating conditions. Additionally, TRUSTEE seeks to promote collaboration with other European Cloud Spaces and initiatives, such as GAIA-X, in order to conduct testing with partners outside of the TRUSTEE consortium. Based on the feedback supplied by the end users throughout the use case definition and requirements elicitation phases conducted under WP2 activities, the specifics of these trials are elaborated and identified, along with the further and more detailed definition of Pilot Use Cases of TRUSTEE, namely Energy, Health, Education, Space, Automotive, and Trusted Multi-disciplinary Data Exchange. Prior to the pilots, a cloud environment provided by FORTH will be used to test the developed technologies and functionalities, relevant to the

respective implementation status of the TRUSTEE Platform and, thus, perform the Pre – Demo stage for each Pilot Phase of the TRUSTEE Pan-European Pilot Campaign. The TRUSTEE technology prototype during the development phase and the system prototype during the integration phase may both be improved by employing this infrastructure as a testbed.

PILOT PHASE PARTICIPANTS

The goal of this section is to describe the participants of each Pilot Phase for each Pilot Use Case of TRUSTEE; that is the Partners that will take on the role of stakeholders and end-users in each case. Stakeholders and end-users will reflect the TRUSTEE Personas, namely Data Provider, Model Provider, Consumer, and Developer. Prior to the beginning of each Pilot Phase the participants will be defined and reported in the respective deliverable.

PILOT PHASE PARTNER ROLES AND RESPONSIBILITIES

As mentioned in the previous section, certain Partners will take on the role of end-users/stakeholders. Additionally, technical Partners developing functionalities to be tested during the Use Case Scenarios of each Pilot Phase and for each Pilot Use Case will also have specific roles in the TRUSTEE Pilot Campaign Plan, as follows:

- Pilot Use Case Leaders, a role already defined in the GA and adopted by the partners leading the TRUSTEE Pilot Use Cases.
- Pilot Phase Leader, a role that will be adopted mostly by HMU as the WP5 Leader.
- Pilot Use Case Rapporteurs, who will report the outcomes stemming from pilot testing for a specific Pilot Use Case, a role that will be mostly adopted by Pilot Use Case Leaders for the respective Pilot Use Cases they are leading.
- Pilot Phase Rapporteurs, who will be responsible for reporting the overall outcomes and results of a specific Pilot Phase by integrating the inputs provided by Pilot Use Case Rapporteurs.
- Technical Leaders of the Use Case Scenarios to be tested, a role that will be adopted by the respective ABB Leaders responsible for the functionalities to be tested in each specific Use Case Scenario of a Pilot Phase.
- Technical Assistants, who are directly involved in the technical development of the functionalities to be tested in the Use Case Scenario or who will be taking part in pilot testing of the various Use Case Scenarios during each Pilot Phase.
- Legal Support, a role that will be adopted mostly by UNIVIE, who will investigate the legal aspects of the Use Case Scenarios to be tested in each Pilot Phase.

By adopting this approach of distributed effort, the aim is to foster collaboration among TRUSTEE partners and ensure the smooth definition, operation, and evaluation of each Pilot Phase and, eventually, of the entire Pilot Campaign.

PILOT PHASE DATA ACQUISITION AND EXCHANGE

This section outlines the methodology of collecting the information about data aquisition and exchange in different Pilot Phases of the TRUSTEE Pan-European Pilot Campaign.

METHODOLOGY

In all phases of the pilots, it is crucial to track and monitor the activities pertaining to different types of data at the pilot sites. Deliverables in WP5 should therefore include detailed information on how and by whom the data being the input and output data of the Pilot Phases will be handled and managed on the ABBs provided in TRUSTEE, in each specific Use Case Scenario.

In each Pilot Phase and environment, a Data Management Handling Plan[3] (hereinafter DMHP) should be prepared. This section provides a template for the DMHP documents, presented in Table 12, Table 13, Table 14,

Table 15, and Table 16 regarding the following directions, respectively: Data Production and Storage; Organization, Documentation, and Metadata; Data Access; Data Sharing and Reuse of Data; Data Preservation and Archiving.

The DMHP is to collect the details, which will assist the Consortium Partners in assessing the privacy and confidentiality of the processes involving data, as well as in ensuring the privacy and confidentiality of the tested ABBs. Furthermore, the DMHP supports the process of establishing the legal and ethical standards for the whole lifecycle of the data, i.e., data generation, collection, storage as well as use and sharing (including providing access to data) via tested / demonstrated ABBs.

The below template takes into account the open data policy of the European Commission and aims to create a FAIR approach to data and make them Findable, Accessible, Interoperable and Reusable, as well as open, where possible. It also takes into account the guidance of the European Commission on Open Science in Horizon Europe, requiring constant monitoring and reporting on the research outputs.

The inputs to the DMHP foreseen for different phases of the pilots will support the provision of the updates to D1.7 – "Data management plan final", due in M24, and further allow for the collection of relevant updates after the final DMP has been reported.

The DMHP should support the project to outline the following:

(a) How project pilot data will be handled in respective pilot phases?

(b) What kind of data will be collected, generated, or processed?

(c) What standards and methodology will be applied in the pilot phases?; and

(d) Whether data will be shared /made open access/ how data will be curated and preserved?

The tables will be filled out at every phase of the pilots by ABBs developers and the Pilot Leaders, in close cooperation.

As the completion of the information requested under the DMHP will be performed during each pilot phase, the outcomes will be presented respectively: for Pilot Phase 1 – in D5.2, for Pilot Phase 2 – in D5.3, for Pilot Phase 3 – in D5.4, for pilot Phase 4 – in D5.5.

Table 12: Data Production and Storage

[3] Based on D6.1 - Piloting Scenarios & Evaluation Plan | Zenodo of the "The Food Safety Market: An SME-powered industrial data platform to boost the competitiveness of European food certification" project (the FSM), co-funded from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 871703.

DATA PRODUCTION AND STORAGE	
Types or categories of data generated/collected	What types of research data are collected, generated, or produced during a specific pilot phase? Who (or which entity) will be responsible for deciding what data is collected or generated?
Personal or non-personal data	Will the data involved in the pilot phase represent personal or non-personal data? What type of non-personal data will be collected at the pilot site in a specific pilot phase? What type of personal data will be collected?
Dummy/fake or real data	Will the data be dummy/fake or real?
Formats of the data	In which format will the data be collected (e.g., CSV, JSON, xls, PDF, ….)?
Reproducibility of data	Please provide the information for validation and reuse of data and indicate if the data are foreseen as open access
Data size	Please provide the information about the estimated size of data provided as input, as well as foreseen size of the data produced.
Software tools for creating/processing /visualising data	Which application/ABB will be tested in your pilot in this specific phase? What aspects/functionalities of the applications will be tested at your specific pilot location in a specific phase? Besides the ABBs indicated as to be tested in your pilot phase, what other software tools will be used for creating/processing/visualising data?
Use of pre-existing data	Will you use pre-existing data? Yes / No / Uncertain. If so, please indicate what pre-existing data will be used.
Data storage and backup strategies	Please indicate what storage and backup strategies will be adopted.
Purpose of data collection	Considering each type of data collected in the pilot phase, what is the purpose of their collection?

Table 13: Organisation, Documentation and Metadata

ORGANISATION, DOCUMENTATION AND METADATA	
Standards for documentation of metadata	What standards will be used for documentation and metadata (e.g., Digital Object Identifiers)? Is there a community standard for metadata sharing/integration?

Best practice/guidelines adopted for data management	Are there any best practices or guidelines, which are foreseen to be applied in the context of organisation and documentation of the data and metadata in the pilot phases?
Tools for formatting data	What type of tools will you use to format data in the pilot phase?
Directory and file naming convention used.	What directory and file naming convention will be used? Will you provide clear version numbers?

Table 14: Data Access

DATA ACCESS	
Risks to data	What main risks to data collected / produced during the pilot phase do you foresee? • Loss or destruction of data • Data breach • Loss of availability • Loss of integrity • Loss of confidentiality • Unauthorised alteration transmission and storage of data. Please provide any other major risks to data collected/produced at pilot sites.
Risk management	Have you prepared a formal risk assessment addressing each of the major risks to data security and potential solutions? If so, please share further information. If no/uncertain, please explain why.
Data access & requirements for access	Are there any concerns regarding access to your data? Yes / No
Correct execution of the data access process	Please indicate a proper process, which someone would need to take to access data collected/generated at the pilot site during the pilot phase, as well as who is responsible for checking the correct execution of the access process. If data is confidential (e.g. personal data not already in the public domain, confidential business information or trade secrets), are there any appropriate security measures in place or any formal standards that you have to comply with?
Procedures to follow in the event of a data breach	Are there any specific data breach procedures, which you foresee should be followed in the case of such an event?

Table 15: Data Sharing and Reuse

DATA SHARING AND REUSE OF DATA	
Organization/labelling of Data for easy identification	How will you organise or label the data to ensure that researchers may easily isolate fields of interest in their study?
Data Sharing & Audience for Data Sharing	Who can access data produced in the pilot and in a specific pilot phase?
Data Sharing Requirements	Are there any data-sharing requirements, which should be followed in the context of sharing the data produced/generated in the pilot in its specific pilot phase?
Re-use of data	Will the data produced or generated in the pilot during its specific phase made reusable or openly accessible? Will the data be reproducible (i.e. able to be copied)?
Audience for re-use	Who will use the data during the pilot? Who will use it afterwards?
Restrictions on the re-use of data	Are there any restrictions regarding the entities that can re-use the data and for what purposes the data can be used?
Publication of data	Do you plan to publish the data generated / collected in the pilot within its specific phase and if so, then where will you publish them?

Table 16: Data Preservation and Archiving

DATA PRESERVATION AND ARCHIVING	
Archiving of data for preservation and long-term access	How will the data produced within the pilot in its specific pilot phase be preserved for long-term access?
Time period for data retention	How long the data should or could be retained?
File formats of retained data	Please provide in what formats the data will be retained.
Data archives	What type of data archives will be used to retain pilot generated/collected data?

Long-term maintenance of data (systems and procedures)	Please provide the details on envisioned systems and procedures for the long-term maintenance of data.

PILOT PHASE LEGAL AND ETHICAL CONSIDERATIONS

Specific legal and socio-ethical aspects relevant for the use case scenario in a specific pilot phase can be identified and addressed only when each of the intended scenarios are defined.

Once it is established which scenarios will be conducted per pilot phase, the Consortium will have to take into account the legal framework and socio-ethical considerations being part of D2.1 [1] and analyse their relevant contextual applicability.

The parties involved in each pilot phase and specific scenarios, with the support from UNIVIE and EPL, will consider if there are any existing additional legal and socio-ethical requirements which may have not yet been indicated in the initial legal framework presented in D2.1 [1], but nevertheless should be addressed with regards to each Pilot Phase.

Additionally, the Pilot Leaders should consider and address the requirements identified for the TRUSTEE Consortium in D1.6 [4], in the context of data protection, for example:

- Compliance with the GDPR in the case of processing real data representing personal data:
  - Identification if the data are personal data.
  - Implementation of the solutions addressing principles of personal data processing
  - Indication of a legal basis for processing personal data
  - Provision of information about data processing to the data subjects
  - Established collaboration between controllers and processors of personal data, as well as recipients and third parties.
  - Data Protection Impact Assessment conducted where required.
  - Maintenance of the records of personal data processing activities
  - Implementation of privacy and security by design approach
  - Applied security measures ensuring appropriate protection of the data of the pilot participants and other persons.
  - Adherence to any local data protection laws
  - Cross-border sharing of the data in compliance with personal data protection laws.

The assessment of the compliance of the data processing activities with Legal and Socio-ethical requirements should be ensured and monitored in each Pilot Phase.

PILOT PHASE EXPECTED OUTCOMES

Prior to the execution of each Pilot Phase, a set of expected and foreseen outcomes of the phase will be presented and discussed with regard to adoption and usage, data economy growth, scalability and integration, security and data privacy, as well as training and knowledge transfer. Alongside the objectives of the respective phase, the expected outcomes will assist in the assessment of the general outcome of each Pilot Phase after its completion.

TRUSTEE PAN-EUROPEAN PILOT CAMPAIGN EVALUATION FRAMEWORK

This section describes the Evaluation Framework that will be used during the TRUSTEE Pan-European Pilot Campaign in order to evaluate and assess the results and outputs of each Pilot Phase as well as of the entire campaign. The aim, research questions to be investigated, methodology for the conduction of each Pilot Phase, data collection and analysis methods and tools to be considered, and the measurement approach to be employed for assessing the realization of the KPIs of each Pilot Phase and the entire campaign are outlined in this part of the document as a procedure and guidelines to be followed after the end of each Pilot Phase.

The above-outlined approach is an initial methodology foreseen to be used during the campaign; however, it may be revised and improved depending on additional demands that may arise during the project and the campaign itself.

AIM

Each Pilot Phase of TRUSTEE's Pan-European Pilot Campaign has a specific goal in the context of the whole campaign, as well as several objectives that are specified in the centre of the Phase and drive pilot testing. Prior to the start of each Pilot Phase, the objectives will be set and clearly defined, as indicated in the Definition Framework, while they may be refined as the respective phase progresses. Then, the aim of each phase will be conveyed while reviewing the findings and outputs of the phase and will be the focal point of the assessment itself.

RESEARCH QUESTIONS

This sub-section presents the Research Questions (RQs) that will be at the core of the TRUSTEE Pilot Campaign and will drive the evaluation process. General RQs have been considered and are outlined below, while any additional Pilot Phase-specific questions will be defined prior to the initiation of each Pilot Phase.

TRUSTEE aims to build a secure-by-design federated data operations platform that will support fair and ethical data collection, transmission, storage, processing, and manipulation in accordance with the principles of responsible/trustworthy AI by using social innovation and a co-development approach as the foundational methodology. This will include assuring compatibility with existing data platforms (such as GAIA-X), enabling cross-border scenarios, and scaling a variety of AI-based applications. The main RQs that TRUSTEE aspires to investigate are stirred around its core objectives and the scientific/technical, social, and economic impacts that it aims to achieve, regarding fostering collaboration among diverse actors in data spaces, such as data providers and consumers; building trusted data solutions with a focus on trustworthiness; and automating testing and monitoring of the health of shared data in a distributed manner in order to enhance traceability and accountability. Some of the general but core RQs of the entire TRUSTEE Pan-European Pilot Campaign include the following:

- RQ1: How is a collaboration between diverse actors fostered within TRUSTEE?
- RQ2: How is European leadership promoted in the global data economy by the solution delivered by TRUSTEE through enabling actors to contribute and glean insights from other services?
- RQ3: How is data trustworthiness extended and widened through the use of the TRUSTEE Platform for multi-disciplinary data use?

- RQ4: Are there benefits created at the social and/or economic level through the delivery of the final prototype of the TRUSTEE Platform?
- RQ5: How is the health and age of shared data monitored within TRUSTEE?
- RQ6: Are there benefits stemming from the use of the TRUSTEE Platform with regard to the management of societal challenges?
- RQ7: How is the user-driven approach adopted for the development of the TRUSTEE Platform answering open legal, socio-ethical, and technical questions in the realm of data spaces regarding multi-disciplinary data users?
- RQ8: What are the benefits offered by TRUSTEE in enabling novice users to use data from multi-disciplinary sources with regard to data interoperability?
- RQ9: How are theoretical estimates of cost-effectiveness approached within TRUSTEE?
- RQ10: How are scientific and technological expertise and know-how enhanced within TRUSTEE?
- RQ11: Are the perspectives of individuals changed regarding how data are/should be used and shared through the use of the TRUSTEE Platform?

The aforementioned RQs will be enhanced and complemented by more specific RQs that will be defined for each Pilot Phase of the campaign and will pave the way towards evaluating and assessing the results and outputs of the respective phase.

LITERATURE REVIEW ON PILOT EVALUATION METHODS

Essentially, piloting involves measuring the baseline level of an identified outcome before running the service or testing the solution, and then measuring the same outcome in the same way after running or testing the solution [11]. This will give a measure of distance travelled and can indicate that the solution or service has the potential to improve the outcomes stated.

Most of the evaluation methods analysed identify the following evaluation cycle to be carried out for a successful and comparable evaluation, as showcased in Figure 4:

1. Identification of the KPI indicators. In several methodologies, this selection is done more systematically, by reviewing standards [12] or based on the requirements of the project itself [13]
2. Select key dimensions to be evaluated for supporting the consecution of the KPI indicators.
3. Select evaluation mechanisms that support the comparison between the "business-as-usual" situation without the result of the trial and the result after the trial.
4. Collect data and monitor the indicators to allow the identification of the benefit obtained.
5. Graphical representation of the results mainly in the different key dimensions and feedback to improve the process.

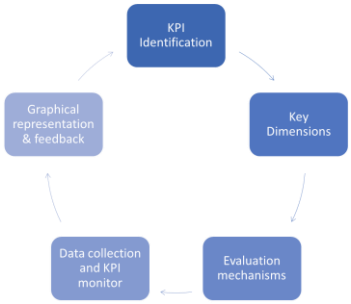

Figure 4: Evaluation cycle based on [3]

These activities are assigned to different phases on the pilot execution [13], [14].

1. Preparatory phase: The main activities here are to define the KPI and Key dimensions of the pilot and identify mechanisms for evaluation.
2. Implementation phase: This phase can be repeated as many times as defined in the project (at least twice). During this phase, moreover the execution of the solution in the pilot, the main activities are to collect the data and monitor the KPIs.
3. Analysis phase: Its objective is to present the results and the improvements achieved. During this phase, it is important to collect the lesson learnt not only from the solution point of view but also from the process point of view.

In the field of pilot evaluation methods, as commented in [12] there are several standards and recommendations that indicate some aspects and KPIs to be taken into account when different systems are evaluated or provide principles to carry out stakeholders' evaluation or surveys for different domains [15], [16]. KPIs can also be identified according to the characteristics and commitments of the project.

Key dimensions [13] should help to understand whether the system evaluated in the pilot project is working well enough to achieve the overall project objectives and the identified KPIs. The most commonly used Key dimensions are the ones outlined in the list below:

- Functionality – Does it work well?
- Reliability – Does it work every time you use it?
- Usability – Is it user-friendly, particularly for novice users?
- Suitability - Is it a good fit for the given context and locality?
- Robustness – Can it operate in the required environment under prevalent conditions?
- Maintainability – Can it be easily maintained?

For each one of the Key dimensions different metrics should be defined (examples could be found [13]) to be able to monitor the progress of the KPI in the different phases of the evaluation. Moreover, to collect data to support the analysis of the different dimensions, it is important to obtain the feedback of the pilots in order to integrate it, if possible, in the following phases of the piloting. The last activity in the preparatory phase is to select the evaluation mechanisms. These evaluation mechanisms support the data collection and the KPI monitoring, during the implementation phases. The evaluation mechanism is selected depending on the type of pilot that is been conducted [14].

There are different evaluation mechanisms depending on the type of data wanted to be collected. For example, to collect qualitative data the following methods can be used: interviews, focus groups, and observations. In the case of quantitative data, surveys and field tests (pre and post-test collecting baseline metrics for comparison) can be employed. Regarding field testing, which is one of the mechanisms used in TRUSTEE Pilot evaluation, it is relevant to select the most appropriate one:

- Experiments, which follow a more rigorous approach: Random Assignment of the participants in two groups one with experience and the others without, establishment of the baseline measures, data collection after the testing to gather the outcome measures and finally statical methods to analyse the results of the experiment.
- Quasi-experiments, which satisfy some steps of the experiment but not all. For example, a non-random assignment or not collecting baseline measures.

Once the preparatory phase with the definition of the KPIs, the dimensions, and the evaluation mechanisms are concluded, the following phases can start following the plan and making as many interactions as stipulated in the evaluation plan.

EVALUATION METHODOLOGY

This section outlines the general methodology to be followed for the evaluation of the results and outputs of each Pilot Phase considering KPIs and their measurement alongside the initial approach towards UX Evaluation of the TRUSTEE Platform. An initial set of methods and tools to be used for data capturing and analysis during and after pilot testing are also discussed below, laying the common ground upon which each Pilot Phase will be approached. The methodology described serves as an initial set of guidelines, which may be revised, enhanced, and complemented depending on additional demands that may arise during the project and the campaign itself.

Methodology refers to the overall research strategy utilised to carry out research. It focuses on the systematic and structured methods performed by researchers to conduct investigations based on the set research goals [17].

Methodology in research combines various approaches such as theoretical and philosophical aspects, research design, data collection methods, data analysis techniques, etc. It provides a structure that defines how research is planned and carried out to help researchers to make appropriate decisions about the suitable methods to be adopted in the research. While providing an outline of how research is conducted, it also specifies the techniques and procedures which to be adopted in order to distinguish information related to a particular research area. The research methodology thus focuses on the way a researcher designs their study enabling them to acquire valid and well-founded outcomes so that the research objectives can be met [18].

In contrast, research methods focus on particular methods, procedures or tools which are leveraged by researchers to obtain, analyse and interpret data. Research methods can be quantitative, comprising numerical data or qualitative, comprising non-numerical data.

Generally, a formal research methodology defines what, by whom, how to collect, and how to analyse the data. It focuses on describing the reasoning behind the adopted approaches in order to justify the key methods of the research. Particularly, it should indicate the methodological choices by justifying why they were selected. In addition, it should establish that the chosen methods are appropriate to acquire definitive and reliable outcomes to support the aims and objectives. In summary, the structured method followed to determine the solution to a problem is Research methodology [19].

There are several focused areas in methodology which include:

- the research method
- the reasonings behind the chosen methodological approach
- the data collection method
- the selected method to analyse the collected data
- familiarise readers with any non-standard approach
- the sampling procedures, and
- limitations.

SIGNIFICANCE OF METHODOLOGY IN RESEARCH

The study of methodology helps to pick the best method, data, and scientific ways and also educates the procedures to solve the problems. To address the research obstacles, it is crucial to have a well-crafted methodology. It should not only focus on recognising the problem but also on establishing the best method that solves the problem. Methodology plays a significant role in solving the problem in research as follows:

- It enables researchers to determine the suitable method for addressing the research challenges.
- It illustrates the effectiveness of the methods in solving the problems.
- It helps to learn the precision of the way decided to apply in research for a satisfactory outcome.

TYPES OF RESEARCH IN RESEARCH METHODOLOGY

Researchers employ a systematic and logical approach to acquiring useful information. Thus, various methods of research are followed that fall under a research methodology, including (1) Basic research (2) Applied Research (3) Problem-oriented research (3) Problem-Solving research (4) Qualitative research and (5) Quantitative research.

TYPES OF DATA IN RESEARCH METHODOLOGY

Data can be raw and unorganised facts that require processing in order to be meaningful and useful. Without proper organisation, the data often becomes of little value. Various data collection methods are employed by researchers, and after data collection, it needs to be processed, organised and structured in a way that turns into useful information. Data can be collected in multiple means and researchers assign particular values to the acquired data. Different types of data in research methodology can include (1) qualitative data (2) quantitative data (3) categorical data (4) observational data (5) experimental data (6) simulation data (7) derived/ compiled data.

USER EXPERIENCE EVALUATION

GENERAL APPROACH AND GUIDELINES

UX refers to the overall experience and satisfaction that the end user has when interacting with a product, system, or service. The evaluation of the experience aims to assess the overall proposed value of the TRUSTEE Platform in terms of the user's interaction, including their perceptions, emotions, and behaviours. Through a variety of widely utilized means to evaluate UX and TRUSTEE-specific surveys, monitorable KPIs are expected as outcomes. Through these KPIs, the end-user experience in the TRUSTEE Platform can be monitored improved and validated.

UX EVALUATION PLAN

The evaluation plan outlines the approach and procedures for assessing the end user experience in the TRUSTEE Pilots. It essentially serves as a roadmap to conduct the evaluation, gather valuable insights, and identify areas for improvement. A UX evaluation plan commonly foresees the following steps [20]:

- **Scope Definition:** the scope of the evaluation should be clearly defined, including the name of the Pilot and an approximation of how much of the pilot the test will cover (e.g., The navigation; the different ABBs; the navigation and ABBs)

- **Purpose Definition:** the concerns, questions, and test goals need to be identified. The purpose can remain quite broad, for example, whether the end user can access the TRUSTEE Dashboard under the Space Pilot Use Case
- **Schedule & Location Indication:** the place and time of the UX evaluation test need to be specified. By setting the schedule of the evaluation test, the different sessions under the test can be further defined.
- **Sessions Definition: t**he different sessions under the same test shall be defined. Typically, an hour or an hour and a half sessions are preferred, with gaps of some minutes between sessions to reset the environment and foresee possible participant or previous session delays.
- **Equipment Description: t**he equipment used in the test should also be described (e.g., desktop; laptop; smartphone;) and its technical specifications should be logged. Also, if the session will be recorded, audio taped, or if special accessibility tools will be used shall be defined.
- **Participants Number:** both the number and the type of participants that will participate in the UX evaluation process should be indicated.
- **Scenarios Definition:** the number and nature of tasks to be included in the UX evaluation process should be defined. Typically, a test lasting one (1) hour shall end up with approximately 10 +/-2 scenarios. However, the number of scenarios will be adjusted to the respective Pilot Phase as well as the overall maturity of the TRUSTEE Platform during the UX evaluation.
- **Subjective Metrics Definition:** not to be confused with the Evaluation Metrics described below, Subjective Metrics include questions to be asked before the evaluation session for example participant background questionnaire or overall satisfaction and likelihood to recommend questions past the session.

UX EVALUATION METRICS

Evaluation metrics are quantitative and qualitative measurements providing objective data and insights and, within TRUSTEE, they will be used to evaluate the overall effectiveness and usability of the TRUSTEE Platform.

- **Successful Task Completion Rate:** this metric measures the percentage of users who successfully complete a given task or scenario during the Pilot evaluation.
- **Critical Errors:** a critical error could be considered a deviation from the scenario goal. For example, returning wrong data according to the specified workflow results in the inability of the participant to complete the task.
- **Non-Critical Errors:** non-critical Errors can be considered recoverable deviations from the defined scenario. For example, the participant opens the wrong navigation menu or using a form incorrectly.
- **Error-free rate:** error-free rate is the percentage of participants that complete the given task without any critical or non-critical errors.
- **Time On Task:** time on task is the amount of time a participant needs to complete a given task.
- **Subjective Measures:** subjective Measures are ratings reported by the evaluation participants, on a 5-to-7-point Linkert scale, to measure satisfaction, ease of finding information, ease of use etc.
- **Likes, Dislikes and Recommendations:** participants will define what they like most and/or least about each evaluation test and possible recommendations.

RUNNING THE UX EVALUATION

Once the evaluation plan has been defined each evaluator can begin the UX evaluation process. TRUSTEE consortium partners will be designated to adopt the role of evaluators during a UX evaluation process. A test run is highly recommended before running the actual evaluation. In the test run, the equipment and materials are being tested with a volunteer participant. Such a process allows for equipment testing, provides practice for the evaluator, defines if the questions and scenarios are clear to the participant and finally allows for last minute adjustments.

After running the test sequence, the evaluator can proceed with the UX evaluation process. To improve the procedure, the following best practices could be applied:

- Treat participants with respect
- Make participants feel comfortable.
- Remember that the TRUSTEE Platform is being evaluated, not the user. Help the participants understand that they are helping to evaluate the TRUSTEE Platform.
- Remain neutral. If the participant asks a question, it is advised to reply with "What do you think?" or "What would you do?".
- Do not jump in to help or lead the participants. If participants give up, it is preferable to give a hint instead of interfering.
- Take good and analytical notes. The more detailed the notes are the easier the analysis.
- Keep track of performance and subjective measures (participant likes and dislikes) independently. People's performance and preferences do not necessarily match.

Following the evaluation plan and applying the best practices described above, the UX evaluation of TRUSTEE should fall among the following example:

1. The evaluator welcomes the participant.
2. The test session is explained to the participant.
3. The participant signs the release form.
4. The evaluator explains where to start.
5. The participant reads the task scenario aloud and begins working on the scenario while thinking aloud.
6. The note-takers take notes of possible comments and/or errors and enumerate behaviours and possible task success or failure that the participant conduct.
7. The session continues for each task of the scenarios or when the defined time runs out.
8. The evaluator asks end-of-session questions in person or through an online survey.

After the sessions is over, the evaluator resets the materials and equipment, has a brief discussion with the observers and waits for the next participant to arrive.

REPORTING UX EVALUATION RESULTS

At the end of the UX evaluation, several types of data will have been collected. When analysing the collected information, it is crucial to organize it in terms of the metrics nature. Specifically, the notes taken during the evaluation process are to be taken into consideration in order to identify possible patterns and attempt to further describe possible problems or errors. In Table 17, some suggested observations are being enumerated and categorized with respect to their quantitative or qualitative nature.

Table 17: Quantitative and Qualitative Data to be collected.

Quantitative Data	Qualitative Data
Success rate	Observations about workflows
Task time	Experienced Problems
Critical Errors	Comments
Non-Critical Errors	Recommendations
Satisfaction questionnaire rating	Answers to pre-evaluation questions
Error Rates	Answers to post-evaluation questions

TRUSTEE QUALITY OF EXPERIENCE ASSESSMENT FRAMEWORK

To the best of our knowledge, the research on Perceptual quality metrics and User Acceptance metrics for distributed data sharing platforms is in its infancy. Therefore, and based on the general approach and guidelines outlined above as well as on relevant literature, we are developing a hybrid Quality of Experience (QoE) assessment framework based on relative research on adjacent technological markets (i.e., user acceptance and QoE evaluation of computer systems, software and cloud-edge networks). The objective of the User Acceptance metrics is to determine the acceptability of different kinds of services (in this case we are focusing on large scale data use and re-use platforms). Using [21] we describe acceptability as the "prospective judgement" made by a group of potential users regarding the adoption of a given service or technology, whereas acceptance, refers to the actual adoption behaviour demonstrated by them when the service or technology is available. The assessment will build on:

1. The user-acceptance models proposed by Venkatesh and colleagues [22] that correlate acceptance with the constructs of perceived usefulness and perceived ease-of-use. The Technology Acceptance Model (TAM) is a widely used theoretical framework. TAM aims to explain and predict users' acceptance and adoption of new technologies. It proposes that users' intentions to use technology are primarily influenced by two factors: perceived usefulness (the extent to which a user believes the technology will enhance their performance) and perceived ease of use (the user's perception of how easy it is to use the technology). TAM has been influential in understanding technology adoption and has been applied across various domains.
2. The Unified Theory of Acceptance and Use of Technology (UTAUT) [23], [24] is an extension and integration of several existing technology acceptance models. Introduced by Venkatesh, Morris, Davis, and Davis in 2003, UTAUT aims to provide a comprehensive understanding of technology acceptance and usage behavior. It includes four key determinants: performance expectancy, effort expectancy, social influence, and facilitating conditions. UTAUT considers social and contextual factors, as well as hedonic motivation and price value, to explain user behavior related to technology acceptance.

Figure 5 below shows a graphical depiction of the assessment model. The blocks represent the model constructs while the arrows highlight the known correlations. The evaluation will focus on the Technology Acceptance Model blocks plus the objective usability. Given the privacy and security aspects of a data use/re-use platform accessed by users with diverse technological background, employees, etc., metrics of trust and perceived safety are also determined, as well as the system's usability and user error tolerance. Unless otherwise stated, a psychometric scale composed of a set of questions answered through a Likert scale [22] will be used to assess each identified metric. The complete set of questions addressing all metrics will be contained in a questionnaire provided to end-users, adhering to the common, unified measurement methodology presented in this deliverable. Questionnaires will be typically answered after the scenarios and games. When possible, objectively measured KPIs addressing the system's usability will serve as a complement to the self-assessed results.

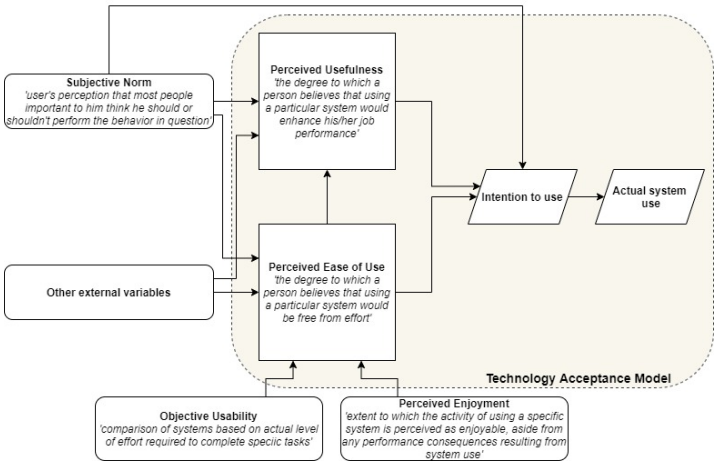

Figure 5: Simplified Technology Acceptance Model to be considered for TRUSTEE Platform user acceptance evaluation.

For each of the metrics evaluated through a psychometric scale, a group of questions/statements will be defined based on pre-validated user-acceptance scales, with adaptations (if required) for each specific TRUSTEE Pilot Use Case. The respondent will answer each question/statement through a 5-point Likert Scale ("Strongly Disagree -> Strongly Agree"). The use of multiple questions per construct allows for a stronger internal validity and reliability of the scale [25].

QUALITY OF EXPERIENCE METRICS

In the scope of the quality of experience and user acceptance evaluation activities, we consider end-users the stakeholders that have direct interaction with the TRUSTEE Platform.

The metrics presented in this section are divided into four different categories. The first one refers to metrics of technology acceptability and is based mostly on the work in [26] and [27]. The second one refers to measures of trust and perceived safety, which are of the essence when referring to data-sharing platforms. The third refers to system usability as measured by observation of the interaction, which is an indicator of acceptability. The last one refers to the ability of the system to deal with user error and misuse. A questionnaire will be created to collect the subjective evaluations of the following QoE KPIs. The participants during the pilot validations will be asked to answer these questions during two iteration periods, namely M27-M32 and M34-M39, which relate to Pilot Phases 3 and 4, respectively. The evaluation of the responses and the statistical analysis will be provided over the final deliverables of WP5.

Technology Acceptability Metrics

- ***QoE1.1 – Acceptance***
  - A psychometric scale rating of the acceptance intention (acceptability) regarding the evaluated use-case. Acceptability is defined as the "prospective judgment" made by potential users regarding their adoption of the system or technology.
- ***QoE1.2 – Perceived Technology Usefulness***
  - A psychometric scale rating of the perceived technology usefulness regarding the evaluated use-case. The perceived technology usefulness is defined as the extent to which the respondent believes that the service/technology will facilitate his/her achievement of a task/goal at hand.
- ***QoE1.3 – Perceived Technology Ease-of-use***
  - A psychometric scale rating of the perceived ease-of-use regarding the evaluated use-case. The perceived ease-of-use is defined as the extent to which the respondent believes that the service/technology is easy to use.
- ***QoE1.4 - Affinity for Technology Interaction***
  - A psychometric scale rating of the user's general ability for interacting with technological artefacts. Several researchers point this factor as relevant in understanding user acceptance. This metric will employ the Affinity for Technology Interaction (ATI) Scale [28].
- ***QoE1.5 - Acceptability difference between prior and post-contact with technology***
  - For the test subjects that interact with the technology, the variation in terms of acceptance intention, perceived usefulness, and perceived ease-of-use between before (prospective evaluation) and after (retrospective evaluation) contact with the technology. The evaluation of this metric will focus on the comparison of the scores of metrics QoE1.1, QoE1.2 and QoE1.3 prior and after contact with the technology.

Unified Theory of Acceptance and Use of Technology (UTAUT) Metrics

- ***QoE 2.1 - Social Influence (SI)***
  - Social influence refers to the degree to which an individual perceives that important others, such as friends, colleagues, or experts, believe they should use a particular technology. It captures the impact of social factors on technology acceptance and usage behaviour. This metric will be compared against QoE1.1, QoE1.2 and QoE1.3 and the analysis can help identify the relative importance of social influence in relation to other factors.
- ***QoE 2.2 - Facilitating Conditions (FC)***
  - Facilitating conditions represent the external factors that can either support or hinder the use of technology. It includes factors such as the availability of technical support, infrastructure, resources, and training that can influence users' ability to adopt and utilize the technology effectively. This metric will be compared against QoE1.1, QoE1.2 and QoE1.3 and the analysis can help assess the relative importance of facilitating conditions compared to other factors in influencing technology acceptance and usage.
- ***QoE 2.3 - Hedonic Motivation (HM)***
  - Hedonic motivation refers to the extent to which individuals perceive that using technology will provide them with enjoyment, pleasure, or fun. It reflects the intrinsic or experiential aspects of technology use that go beyond the purely instrumental or task-oriented benefits. This metric will be compared against QoE1.1, QoE1.2 and QoE1.3 and the analysis can help assess the relative importance of hedonic motivation compared to other factors in influencing technology acceptance and usage.
- ***QoE 2.4 - Price Value (PV)***
  - Price value represents users' perception of the relationship between the cost or price of using technology and the benefits derived from it. It considers the economic or financial aspect of technology adoption and examines how users evaluate the value proposition of the technology in relation to its cost.

Trust and Perceived Security Metrics

- ***QoE3.1 - Perceived Security***
  - A psychometric scale rating of the perceived security of the system evaluated in the user-story. Perceived security is a construct defined as the extent to which an individual believes using the system will carry some risk to his security and privacy of personal data, etc.
- ***QoE3.2 - Perceived Trust***
  - A psychometric scale rating of the perceived trust in the system evaluated in the user-story. Perceived trust is a construct that defines the extent to which the individual believes that the system/technology will assist him in achieving a goal even in uncertain and vulnerable situations.

Systems Usability metrics

- ***QoE4.1-General usability metric***
  - For the test subjects that interact with the technology, a psychometric scale score of the system's perceived usability. This metric will employ the System Usability Scale (SUS) [29]. This is ten items scale with questions such as: *"I think that I would like to use this system frequently"*; *"I found the system unnecessarily complex"*.
- ***QoE4.2-Effectiveness***
  - For the test subjects that interact with the technology, a score of the system's effectiveness (i.e., level of success) in handling the human-machine interaction. This metric will be assessed based on: (i) Percentage of sub-tasks (within each task) achieved (where applicable); (ii) Percentage of users successfully completing the task. This metric shall be assessed through the means system's event log data, where applicable (apart from the questionnaire). The goals will be defined per use case, based on the human/machine interactions that are expected to be conducted. The contribution of each metric to the final overall score will be determined based on system analysis by technology experts.
- ***QoE4.3-Efficiency***
  - For the test subjects that interact with the technology, a score of the system's performance level in handling the human-machine interaction will be assessed on the basis of the following (second-level) metrics: (i) Time to complete the task; (ii) Number of instances user diverted from the scenario path (where applicable); (iii) Psychometric scale for Mental Workload – using the Nasa Task Load Index (TLX) questionnaire [30]. Values shall be assessed through the system's event log data, where applicable (apart from the questionnaire). The contribution of each metric to the final overall score will be determined based on system analysis by technology experts.
- ***QoE4.4-Satisfaction***
  - For the test subjects that interact with the technology, a score of their satisfaction in their interaction with the TRUSTEE Platform. This will be assessed on the basis of the following (second level) metrics: (i) Psychometric scale for satisfaction – using the After-Scenario Questionnaire (ASQ) [31]; (ii) Frequency of complaints (10% or less dissatisfaction); (iii) Psychometric scale for the feeling of frustration using NASA_TLX [30]. Values shall be assessed through the system's event log data, where applicable (apart from the questionnaire). The contribution of each metric to the final overall score will be determined based on system analysis by technology experts.

Error Tolerance Metrics

This section lists KPIs to evaluate the TRUSTEE Platform's ability to deal with user error and misuse. These will be applicable for pilot use cases that imply an interaction between users and the TRUSTEE Platform.

- ***QoE 5.1-Error dealing effectiveness***

  - For the test subjects that interact with the platform, a score of the system's effectiveness to deal with user errors. This will be assessed based on the following (second level) metrics: (i) Percentage of errors reported by the system; (ii) Percentage of user errors tolerated. This metric shall be assessed through the system's event log data, where applicable (apart from the questionnaire). The contribution of each metric to the final overall score will be determined based on system analysis by experts.

- ***QoE 5.2-Error dealing efficiency***
  - For the test subjects that interact with the platform, the percentage of time spent on correcting interaction errors. This metric shall be assessed through means of observation (Video), where applicable.
- ***QoE 5.3-Error dealing satisfaction***

For the test subjects that interact with the platform, a psychometric scale rating of their satisfaction with the system's ability in dealing with user errors [31].

DATA COLLECTION FOR IMPACT ASSESSMENT

Besides the KPIs defined at the system level, KPIs for each subsystem depicted in the TRUSTEE Platform breakdown architecture need to be identified. Individual subsystem KPIs will be evaluated and measured as part of the validation scenarios.

To assess the outcome of each validation scenario, the relevant assessment indicators for each component shall include:

- Comparison of outputs
- Measure of time
- Questionnaires to collect users and/or stakeholders' perception.
- Individual interviews
- Group interviews/discussion

QUANTITATIVE CAPTURING METHODS AND TOOLS

Quantitative methods shall include metrics and questionnaires. Questionnaires targeting the several metrics mentioned above regarding UX evaluation will be developed. Pre-validation questionnaires could be used to assess the proposed solution prior to pilot testing. After the deployment of the solutions, a post-validation questionnaire will be filled out in order to collect values. Such questionnaires could have both quantitative and qualitative nature. Additionally, specific metrics for each subsystem could be defined in order to further identify the values that need to be collected, respectively.

QUALITATIVE CAPTURING METHODS AND TOOLS

The goal of qualitative data collection is to identify the types of data that will address the research questions [32]. The principle is to pose general and broad questions to participants and allow them to share their unrestricted views, allowing one to gather a variety of data while adding new types of information as the study progresses to help them with the conducted research questions.

Technically, the qualitative data collection is non-numerical but more textual oriented which includes images, written texts, and recorded audio. It aims at examining the reason for a situation, or phenomena and understanding the experience of people. Qualitative data collection focuses on answering "how and why" questions in a research study.

In the literature, there are five types of qualitative data collection, namely Observation, Depth Interviews, Storytelling, Document Analysis, and Brainstorming, which are further discussed in the following subsections.

OBSERVATION

Observation is a technique that involves watching, and recording the characteristics and behaviour of people, objects, or phenomena, etc. The goal is to increase the sensitivity of each detail and at the same time be able to focus on people, objects or phenomena that are of genuine interest to the study. The observation technique gathers open-ended information. Technically, observations are unstructured texts and pictures taken during observations by the researcher.

DEPTH INTERVIEWS

Depth Interviews are one of the most popular and reliable techniques for gathering qualitative data. This collection technique consists of direct communication with a single person, a group of people, and face-to-face interaction between people. The researcher creates an interview questionnaire to obtain information about the interviewee's knowledge or perspective of a subject, problem, or ideas. Depending on the beliefs, experiences, and point of views of each person, the inquiries in this situation may be more open-ended, structured, unstructured, or informal. Technically, depth Interviews are unstructured texts obtained from transcribing audiotapes of interviews or by transcribing open-ended responses to questions on questionnaires. Interviews are a part of a social interaction and are performed with an individual, or group a group of people [33]. Two methods can be employed to conduct an interview [34]–[36].

- Group interviews focus on a debate within a group by collecting general descriptions. Group interviews are used when it requires interactions and dialogues between individuals. For example, the interaction within a subgroup can express differences of opinion on a topic.

STORYTELLING

Storytelling is a narrative description of events and life experiences by a person who experienced them [37]. Stories could be public and private records available to the researcher, such as notes from meetings and journals. In contrast to interviews, fewer questions serve as a guide for the dialogue, and typically there is just one important question that aids in establishing the description's overall structure.

Within the storytelling technique, three dimensions are examined [38].

- Interaction: Examining both personal feelings and social points of view.
- Continuity: Examining past actions and their impact on current and future experiences.
- Situation/Place: Examining physical environment and setting.

DOCUMENT ANALYSIS

Document analysis consists of gathering information from already existing sources. The data is gathered from personal documents, as well as other sources of information such as scientific literature, and newspapers. Occasionally, documents contain images or sounds of people, life histories, or objects recorded by the researcher or someone else.

BRAINSTORMING

Brainstorming is a group activity-centered technique, in which individuals interact with each other by finding ideas, solutions, and collecting thoughts related to the research topic. It is a group formatted problem-solving and creativity technique addressing a specific question. The objective is to eliminate any critical observation or pressure and generate topics that will be evaluated in each research question. Brainstorming sessions should be created and used at the beginning of a project. There is no limitation on the duration of the session. The session can be short or long.

Table 18 shows an example of the types of data for the five techniques.

Table 18 A Compendium of Data Collection Approaches in Qualitative Research based on [39], [40].

Category	Type of data	Example
Observation	Fieldnotes and drawings	Gather fieldnotes by: • Conducting an observation as a participant • Conducting an observation as an observer • First observing as an outsider then participating in the setting and observing as an insider
Interviews and Questionnaires (Surveys)	Transcriptions of open-ended interviews	• Conduct an unstructured, semi-structured, open-ended interview while taking interview notes. • Conduct focus group interviews • Collect open-ended responses
Document analysis	Hand-recorded notes about documents, pictures, photographs, videotapes, objects, sounds	• Have a participant keep a journal during the research study. • Collect personal notes from participants. • Analyze public documents such as official memos, minutes of meetings. • Analyze school documents such as attendance reports, and discipline referrals. • Examine autobiographies and biographies.
Storytelling	Narrative, statistical analysis, imagery	• Use visuals to show ideas • Know the audience • Outline the core message • Go deep
Brainstorming	Graphic organizers (story map), Visualization, unstructured data	• Identify a potential problem to be solved • Generating, collecting, monitoring ideas • Propose ideas without any self-censorship

A SUMMARY OF DIFFERENCES

Table 19 shows the benefits and drawbacks of each qualitative method. Each method has its own strong and weak points and could be combined for any research topic. The qualitative data collection method is a method used for extracting insights from the data and identifying the behavioural pattern of thinking. In addition, it enables the study of any issue and situation. To strengthen the reliability and integrity of the data collection, qualitative data collection is used to conduct data analysis. There are various ways to acquire qualitative data, and the existing methods are utilized to strengthen the credibility of any result.

Various and complementary methods might be used for gathering no overlapping data for the same topic. Thus, achieving conclusions through different methodologies increases their reliability.

Table 19 Advantages and disadvantages of qualitative methods

Method	Advantage	Disadvantage
Observations	• Collecting direct information. • Involvement of evaluators. • Large samples used in the studies allow for generalization. • High-quality and accurate data can be obtained.	• May be time-consuming • Training of evaluators is effective • Data can be distorted by observers. • Due to the inaccurate representation of the qualitative data measurement, it can occasionally be unreliable.
Interviews	• Collecting rich, in-depth, and detailed data directly • Obtaining knowledge about the past and future for events and features • The flexibility of administration of interviews • Facilitating communication by providing further explanations to questions and answers.	• Hiring and training interviewers • Complex process • Scheduling where and when to meet people and the possibility of changing plans at the last minute • Possibility of missing information • Difficulties and time-consuming of the coding process • Being expensive
Documents analysis	• Less time-consuming • More efficient • Document availability • Cost-effectiveness	• Insufficient details • Retrieving documentation is sometimes difficult or impossible • Biased selectivity [41]
Brainstorming	• Diversity of thought • Quick idea generation • Promote the creativity	• High costs. More expensive than questionnaires. • Time consuming and sometimes hampered innovation. • Fear of judgment and inauthenticity

Table 20 and Figure 6 below show a comparison between qualitative and quantitative approaches. Each approach offers intrinsic methods for data collection processes and they are often combined. For instance, quantitative surveys can include open ended questions for extracting qualitative responses. On the other hand, the qualitative responses can be quantified. Therefore, qualitative and quantitative methods can complement each other.

Table 20: Differences between quantitative and qualitative approaches [42]

	Qualitative	Quantitative
Type of knowledge	Subjective	Objective
Aim	Exploratory and observational	Generatable and testing
Characteristics	Flexible. Contextual based. Dynamic, continuous view of change.	Fixed and controlled. Independent and dependent variables.
Sampling	Purposeful	Random
Data collection	Semi-structured or unstructured	Structured

Nature of data	Narratives, description.	Numbers, statistics
Analysis	Thematic	Statistical

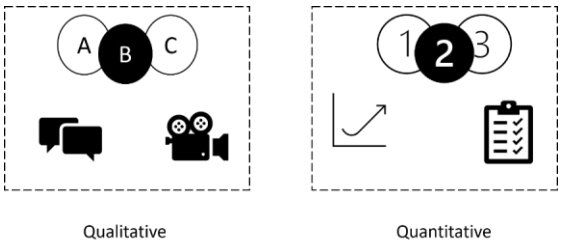

Figure 6 A comparison between the qualitative and quantitative approach

DATA ANALYSIS

This chapter aims to provide a comprehensive overview of the different methodologies to be put in place for the analysis of the data that will be collected during the Pilot Campaign of TRUSTEE and the various Pilot Phases. Different methodologies will be employed to accommodate the different aspects and characteristics of the data collection methods described further above.

By employing these methodologies, both quantitative and qualitative, we can gather comprehensive and diverse data, enabling a holistic understanding of the TRUSTEE system, its subsystems, user experiences, and areas for improvement. The combined use of these methodologies facilitates evidence-based decision-making, supports user-centered design, and enhances the overall effectiveness and impact of the TRUSTEE Pilot Campaign.

ANALYSIS OF DATA CAPTURED BY QUANTITATIVE METHODS AND TOOLS

The use of statistical analysis and data visualization techniques is essential for analyzing quantitative data captured through metrics and questionnaires. Statistical analysis enables the identification of patterns, relationships, and significance within the data, providing objective insights into the performance of the TRUSTEE system and its subsystems. Data visualization complements statistical analysis by visually representing the data, making it easier to comprehend and identify trends or anomalies. Together, these methodologies facilitate data-driven decision-making and enhance the understanding of quantitative aspects of the TRUSTEE Pilot Campaign.

Statistical Analysis: Statistical analysis involves applying mathematical and statistical techniques to analyze quantitative data, providing insights into patterns, relationships, and significance. Descriptive statistics summarize the data, while inferential statistics assess relationships or differences between variables. Correlation analysis determines the strength and direction of relationships, and regression analysis models the relationship between variables.

Data Visualization: Data visualization is the graphical representation of quantitative data to aid understanding and interpretation. It utilizes charts, graphs, and plots to depict patterns, trends, and distributions. Trend analysis visually tracks metric changes over time, comparative analysis compares metrics across different dimensions, and geographic analysis maps spatial variations in metrics.

Comparative analysis and benchmarking methodologies are crucial for evaluating the performance of each subsystem within the TRUSTEE system. By comparing subsystem metrics, researchers can identify variations and performance differences, enabling targeted improvements and optimizations. Benchmarking against industry standards or best practices provides a benchmark for performance assessment and helps identify areas for improvement and potential innovations. These methodologies support evidence-based decision-making and assist in maximizing the effectiveness and efficiency of each subsystem.

Comparative Analysis: Comparative analysis involves comparing subsystem metrics to assess variations and performance differences. Control charts monitor subsystem metrics over time, identifying out-of-control conditions. Pareto analysis prioritizes improvement efforts by identifying the most significant factors contributing to variations.

Benchmarking: Benchmarking compares subsystem metrics against established standards or benchmarks to evaluate performance. It involves comparing performance indicators to industry or sector-specific benchmarks and best practice analysis, learning from high-performing subsystems or organizations.

PRE-VALIDATION QUESTIONNAIRE

Descriptive analysis and group comparisons are used to analyse the pre-validation questionnaire responses. Descriptive analysis allows for summarizing participant characteristics, expectations, and perceptions, providing a clear overview of the user base. Group comparisons help identify any variations in expectations or perceptions among different user groups, allowing researchers to tailor the system and its implementations based on user segmentation. These methodologies facilitate understanding user perspectives, establishing baselines for evaluation, and informing the design and implementation of the TRUSTEE system.

Descriptive Analysis: Descriptive analysis summarizes and describes the characteristics, expectations, and perceptions captured in the pre-validation questionnaire. It includes calculating frequencies and percentages for participant responses and examining summary statistics.

Group Comparisons: Group comparisons assess differences in expectations or perceptions among different participant groups. Independent samples t-tests determine statistically significant differences between two groups, while analysis of variance (ANOVA) assesses variations across multiple groups.

POST-VALIDATION QUESTIONNAIRE

Satisfaction analysis and usability evaluation methodologies are vital for assessing the post-validation questionnaire data. Satisfaction analysis enables the measurement of overall satisfaction levels and identification of key drivers of satisfaction. By evaluating user satisfaction across different system components or user groups, researchers can pinpoint areas of improvement and prioritize enhancements. Usability evaluation methodologies, such as SUS scores and task completion rates, provide insights into the perceived usability of the TRUSTEE system, helping identify usability challenges and areas that require refinement.

Satisfaction Analysis: Satisfaction analysis involves analysing responses from the post-validation questionnaire to assess overall satisfaction levels. It includes calculating mean or median satisfaction scores and examining variations across system components or user groups. Importance-performance

analysis evaluates satisfaction in relation to the importance participants attribute to specific system features.

Usability Evaluation: Usability evaluation focuses on assessing the perceived usability of the TRUSTEE system. The metrics reported to be assessed based on TRUSTEE's QoE Assessment Framework will be analysed alongside task completion rates to evaluate how successfully users can accomplish tasks and identify areas for improvement.

ANALYSIS OF DATA CAPTURED BY QUALITATIVE METHODS AND TOOLS

OBSERVATION

Thematic analysis of observational data allows researchers to gain in-depth insights into user behaviours, interactions, and experiences. Through observation, researchers can capture contextual information that quantitative measures may not fully capture. Thematic analysis helps identify recurring themes and patterns, providing a deeper understanding of how users interact with the TRUSTEE system in real-world scenarios and informing improvements or optimizations.

Thematic Analysis: Thematic analysis is a qualitative data analysis method that involves identifying and analysing recurring themes or patterns in observational data. It includes open coding, where initial codes are generated, axial coding to analyse relationships between codes, and selective coding to refine themes based on research objectives.

DEPTH INTERVIEWS

Content analysis of depth interview transcripts allows for rich qualitative insights into participants' experiences, opinions, and suggestions regarding the TRUSTEE system. Depth interviews provide an opportunity for participants to express their thoughts openly, providing nuanced and detailed information that quantitative methods may not capture. Content analysis helps identify emerging themes, underlying motivations, and valuable suggestions, enabling researchers to understand user perspectives on a deeper level and refine the TRUSTEE system accordingly.

Content Analysis: Content analysis is a systematic approach to analyse qualitative data, such as interview transcripts. Inductive coding involves identifying patterns and themes without preconceived categories, while deductive coding analyses data using predefined categories based on research objectives or prior knowledge.

STORYTELLING

Narrative analysis of user stories allows researchers to uncover emotional expressions, challenges, and successes related to the TRUSTEE system. Storytelling provides a holistic view of user experiences and offers insights that quantitative methods may overlook. Analyzing narratives enables researchers to identify common themes, emotional impact, and key elements that contribute to user satisfaction or frustrations. This methodology adds a human-centric perspective, enhancing the understanding of the TRUSTEE system's impact on users' lives.

Narrative Analysis: Narrative analysis examines the structure, emotions, and themes within stories shared by participants. Structural analysis focuses on story elements such as setting, characters, and plot. Emotional analysis explores the emotions expressed in the narratives, and theme identification uncovers overarching themes or motifs.

DOCUMENT ANALYSIS

Text mining and analysis of relevant documents, such as reports, logs, or user-generated content, provide additional insights into the TRUSTEE system. Text mining techniques enable the extraction of valuable information, sentiments, and patterns from unstructured text data. Document analysis helps identify user feedback, system performance issues, or unexpected patterns that may have emerged during the Pilot Campaign. This methodology aids in uncovering valuable insights that complement other qualitative and quantitative methods.

Text Mining: Text mining is a technique used to analyse unstructured text data. Sentiment analysis determines the sentiment expressed in documents, while topic modelling identifies key topics or themes. These techniques enable insights to be extracted from documents such as reports, logs, or user-generated content.

BRAINSTORMING

Brainstorming sessions provide a collaborative approach to generating innovative ideas and potential improvements for the TRUSTEE system. By analyzing the ideas generated during brainstorming, researchers can identify common themes, prioritize suggestions, and explore new possibilities. This methodology leverages collective intelligence, involving stakeholders' expertise and perspectives to drive continuous enhancement and optimization of the TRUSTEE system.

Idea Generation: Idea generation during brainstorming involves generating and categorizing ideas. Affinity diagrams group ideas based on similarities or common themes, and prioritization matrices evaluate and rank ideas based on predetermined criteria such as feasibility, impact, or user value.

These formal descriptions provide a clearer understanding of each data analysis methodology and its specific purpose within the context of the TRUSTEE Pilot Campaign.

MEASUREMENTS

For assessing the achievement of the various KPIs (i.e., general project KPIs, Pilot Use KPIs, Pilot Phase KPIs) a template has been developed in order to provide a roadmap towards the required measurements. Table 21 presents an initial version of the template that will be used to define the different measurements for the achievement of the KPIs, which will be revised and further enhanced and/or updated based on the demands emerging during pilot testing throughout the TRUSTEE Pan-European Pilot Campaign. The following fields will be considered for each KPI:

- **KPI ID**: A unique identifier will be assigned to each KPI.
- **Name**: Each KPI will have also a relevant name.
- **Definition**: This field will provide a clear and precise explanation of the KPI. It will define what the KPI assesses and the scope of its measurement. Finally, this field will also include and define the individual terms that compose the KPI.
- **Calculation Method**: This field refers to a particular process or formula that will be used to calculate the KPI. The calculation method might involve simple arithmetic operations, complex statistical calculations, or sometimes a composite of several different measures. The calculation method might involve simple arithmetic operations, complex statistical calculations, or sometimes a composite of several different measures.
- **Data Source**: This field will provide the origin of the data that will be used to calculate the KPI.

- **Target Value**: This field will define a target value for the KPI in the context of TRUSTEE, considering existing state-of-the-art solutions and literature.
- **Reporting Frequency**: This will define how often this KPI will be calculated and reported.
- **Reference**: This field will provide a scientific reference about the use and the target value of this KPI.
- **Relevance with User Requirements**: This field will enumerate the user requirements that are related to this KPI.
- **Relevance with Technical Requirements**: This field will enumerate the technical/system requirements that are related to this KPI.
- **Verification**: This field will describe in detail how this KPI will be measured and validated in the context of TRUSTEE, considering the operational environments of each pilot.
- **Priority**: This field shows the importance of this KPI in terms of three qualitative values: "High", "Medium" and "Low".

Table 21: KPI and Measurements Template

KPI #ID - Name	
Definition	*<Provide a clear, concise definition of the KPI.>*
Calculation Method	*<Describe how to calculate the KPI. Include formulas if necessary. The method should be clear and repeatable.>*
Data Source	*< Detail where the data for the KPI comes from. This could be a specific system, department, document, or combination of sources.>*
Target Value	*<Define what a successful result look like. his could be a specific number, a range, or a percentage increase or decrease.>*
Reporting Frequency	*<Specify how often this KPI should be reviewed, e.g., daily, weekly, monthly, quarterly, annually.>*
Reference	*<Scientific reference about the use and target values (baseline) of KPI>*
Relevance with User Requirements	*<Enumerate the user requirements defined in D2.1.>*
Relevance with Technical/System Requirements	*<Enumerate the technical/system requirements defined in D2.1.>*
Verification	*<Establish a process for reviewing and evaluating the KPI's effectiveness, considering the operational environment of each pilot and the corresponding scenarios>*
Priority	*<High, Medium, Low>*

DEFINITION OF PILOT PHASE 1: DRY RUN SCENARIO

SUMMARY

This chapter describes the 1st Pilot Phase of the TRUSTEE Pan-European Pilot Campaign and defines the Use Case Scenarios to be tested, the Demonstrators associated, and the KPIs that will be used to evaluate the performance of the Pilot Phase after it has finished. Additionally, the Participants and the Partner Roles and Responsibilities throughout the Pilot Phase are provided in this chapter, alongside an

investigation towards acquisition and exchange of the data that will be produced, generated, or shared, any additional legal and ethical considerations that need to be taken into account, and expected outcomes.

OBJECTIVES

In general, the 1st Pilot Phase refers to the Dry Run of the various ABBs of the TRUSTEE Platform with regard to the functionalities that have been implemented by the beginning of the phase. Since this Pilot Phase is a Dry Run scenario, intercommunication between the various ABBs is not foreseen in this phase, rather it is anticipated to be performed in later Pilot Phases, following the maturity of the ABBs. Therefore, data sharing and data generation will be limited in this first phase. However, investigation towards integration and intercommunication of and between ABBs to be realized in later Pilot Phases has already been initiated to ensure smooth operation of the entire TRUSTEE Pan-European Pilot Campaign.

The main objective of the 1st Pilot Phase is the fast PoC implementation and dry run pilot testing of the initial set of functionalities developed within the various ABBs of the TRUSTEE Platform, which will be demonstrated through mock-up prototypes of the various subsystems, following the incremental deployment strategy.

USE CASE SCENARIOS

The Use Case Scenarios defined for the 1st Pilot Phase are stirred around the functionalities that are implemented by the beginning of the phase and are not specific to the TRUSTEE Pilot Use Cases. Such specific and more complex scenarios will be tested in the next Pilot Phases, following the maturity of the technical development of the various ABBs and the anticipated intercommunication among them. Use Case Scenarios of the 1st Pilot Phase of the TRUSTEE Pan-European Pilot Campaign focus on PoC implementations and mock-up prototypes to assist partners leading the TRUSTEE Pilot Use Cases in familiarizing with the technologies and solutions developed within TRUSTEE by the various ABBs.

Thirty (30) Dry Run Use Case Scenarios are developed for Pilot Phase 1 by using the template defined in the “TRUSTEE Pan-European Pilot Campaign Definition Framework” chapter of this document. These Use Case Scenarios are presented, in a synthetic and condensed way, in the tables included in the following sections below. All ABBs of the TRUSTEE Platform are involved in the defined Dry Run Use Case Scenarios and are the following (more details for the ABBs and the architecture of the TRUSTEE Platform can be found in D2.1 [1]):

- Homomorphic Enabled Data Fusion (HEDF)
- AI Models as a Service (AIMaaS)
- Trustworthy AI Support Design Framework (TAI-SDF)
- Accountable Transactions Recorder (ATR)
- Knowledge Repository (KR)
- Authentication Manager (AM)
- Data Privacy Impact Assessment (DPIA)
- TRUSTEE Dashboard (DA)
- Self Sovereign Identity (SSI) Homomorphic Capable Framework (SSI-HE)
- One-Stop-Shop (OneSS)
- Data Lake (DL)
- Security and Trust Manager (STM)

UCS-001

Table 22: Use Case Scenario 001

Use Case Scenario Description	
Scenario ID	UCS-001
Title	Storing Public Keys in the Data Lake
Scenario Description	This scenario will showcase the usage of APIs to retrieve HE public keys and store them in the PostgreSQL created database within Data Lake.
Digital Solutions/Functionalities to be tested	HE Public Keys storage API POST/GET communication
Technical partners involved	ACCELI
ABBs and/or inner modules of ABBs that are related	DL
User-centered Requirements	USR-003
Technical Requirements	Req-DL-FUNC-2
Legal Requirements	LEG-USR-024
Socio-ethical Requirements	SOC-ETH-TSD-025
Piloting Summary	In this scenario, the Data Lake gets through an API the HE public keys that have been generated from SSI-HE and provided (for the dry run the public keys have been provided via email). Once the public keys are retrieved, the Data Lake stores each public key in the database.
Main Persona(s) involved in the scenario	Data Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	No comments

UCS-002

Table 23: Use Case Scenario 002

Use Case Scenario Description	
Scenario ID	UCS-002
Title	Querying semantically annotated data concepts
Scenario Description	This scenario will showcase the communication of the Data Lake with the GraphDB RDF store that has been set up within Data Lake, for the identification of which data concepts are relevant for the Data Consumer's objectives in using the TRUSTEE Platform. The Data Lake will search the semantic concepts stored. The query will search for the semantically annotated data concepts that have been mapped with a specific domain and as relevant to a specific HE operation.
Digital Solutions/Functionalities to be tested	Semantic concepts retrieval
Technical partners involved	ACCELI
ABBs and/or inner modules of ABBs that are related	DL
User-centered Requirements	USR-001, USR-013, USR-016, USR-022, USR-023, USR-028, USR-029, USR-032, USR-036, USR-060
Technical Requirements	Req-DL-FUNC-1
Legal Requirements	LEG-TSD-069, LEG-TSD-073, LEG-TSD-074, LEG-TSD-082
Socio-ethical Requirements	SOC-ETH-TSD-025
Piloting Summary	In this scenario, the Data Lake requests from the GraphDB RDF store the data concepts that are mapped with the "health" domain and with the

	HE operation "temperature test", which will search for data that have temperature over > 39 degrees Celsius. For this example, this HE operation has been mapped in the GraphDB to the health dataset "eCRF_Rapid_Covid19_Module2_2A_Vitals", so the Data Lake will return the semantically annotated column "Temperature"[4].
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	UCSC (Health dataset), ISS (Space dataset), EPL (Education dataset), ATHENA (Automotive dataset)
Comments / Open issues	The main anticipated outcome of this scenario is to demonstrate the way that the TRUSTEE Platform can inform the Data Consumer on the exact types of data available for their needs and perform this in a semantic way to allow data interoperability.

UCS-003

Table 24: Use Case Scenario 003

Use Case Scenario Description	
Scenario ID	UCS-003
Title	Achieving semantic interoperability for the multidisciplinary pilot using Health and Education data.
Scenario Description	In this scenario, the Data Lake can query the GraphDB RDF data store and retrieve the semantically annotated concepts to be used for any requested operation.
Digital Solutions/Functionalities to be tested	Semantic concepts retrieval
Technical partners involved	ACCELI
ABBs and/or inner modules of ABBs that are related	DL
User-centered Requirements	USR-001, USR-013, USR-016, USR-022, USR-023, USR-028, USR-029, USR-032, USR-036, USR-060
Technical Requirements	Req-DL-FUNC-1
Legal Requirements	LEG-TSD-069, LEG-TSD-073, LEG-TSD-074, LEG-TSD-082
Socio-ethical Requirements	SOC-ETH-TSD-025
Piloting Summary	In this scenario, the Data Lake looks up the GraphDB store in order to identify which concepts are to be used for the HE operation "find students when one of her/his parents has tested positive for COVID-19". In this case, the Data Lake will search and fetch concepts from the Health Dataset (HE version of "person.csv") and the Education Dataset (HE version of "sample school data_Grade9.csv").
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	UCSC (Health dataset), EPL (Education dataset)
Comments / Open issues	The main anticipated outcome of this scenario is to demonstrate the way that the TRUSTEE Platform extracts machine-readable information, allowing interoperability of knowledge (i.e., the machine understands that the concept "Person ID" of the "person.csv" dataset is the same thing/has the same meaning as the "Covid Status" column of the sample school data_Grade9.csv" dataset.

[4] http://purl.bioontology.org/ontology/HL7/C1550577

UCS-004

Table 25: Use Case Scenario 004

Use Case Scenario Description	
Scenario ID	UCS-004
Title	Performing FL
Scenario Description	This scenario will validate the utilization of Federated Learning using the TRUSTEE Platform. The aim is to test the involved functionalities of HE-FL including setting up a group to collaborate via FL, tracking FL iterations and enabling the aggregation of local models, and different FL configurations (personalized, non-IID, constant vs variable aggregation weights).
Digital Solutions/Functionalities to be tested	AIMaaSS:FederatedLearning
Technical partners involved	ATHENA
ABBs and/or inner modules of ABBs that are related	AIMaaS:FederatedLearningModule
User-centered Requirements	USR-081, USR-085, USR-089
Technical Requirements	Req-AIMaaS-FUNC-2, Req-AIMaaS-FUNC-5
Legal Requirements	LEG-TSD-003-006, LEG-TSD-012, LEG-USR-024, LEG-TSD-107, LEG-USR-026, LEG-TSD-092-93, LEG-TSD-095 LEG-USR-016, LEG-TSD-103 LEG-USR-023, LEG-USR-037-038, LEG-TSD-013-014
Socio-ethical Requirements	SOC-ETHTSD-016, SOC-ETH-TSD-011, SOC-ETH-TSD-012, SOC-ETH-TSD-014, SOC-ETH-TSD-018
Piloting Summary	The dry run scenario of FL functionalities in the automotive pilot use-case. This scenario is built on the in-house CARLA-ROS-based environment of ATHENA. The scenario will demonstrate the impact on the performance of AI models used in various automotive tasks when the TRUSTEE Platform is employed in a number of FL rounds.
Main Persona(s) involved in the scenario	Model Provider, Consumer
Foreseen Pilot Assistant	ATHENA (Automotive dataset – demonstration on in-house simulator), PASEU (Automotive dataset – demonstration in real car)
Comments / Open issues	The main anticipated outcome of this scenario is to demonstrate the positive impact on the performance of training AI models via the FL paradigm without any actual loss originating from executing computations in the homomorphically encrypted domain. This scenario will be tested in all pilot phases by incrementally incorporating the functionality of other ABBs. For example, setting up a group of users to collaborate via FL will employ the dashboard and the AIMaaS. Another example is the aggregation of model parameters in the HE domain which will be performed by the HEDF ABB. Moreover, as the phases progress, the focus will shift from open datasets/in-house CARLA ROS framework to real car scenarios.

UCS-005

Table 26: Use Case Scenario 005

Use Case Scenario Description	
Scenario ID	UCS-005
Title	XAI-By design AI models

Scenario Description	This scenario will validate the utilization of the XAI-By design functionality of the TRUSTEE Platform. The aim is to test the employment of deep unrolling techniques (designed and implemented by ATHENA) for producing interpretable by design and efficient in terms of the number of parameters models during the training phase.
Digital Solutions/Functionalities to be tested	AIMaaS:XAI-By design
Technical partners involved	ATHENA
ABBs and/or inner modules of ABBs that are related	AIMaaS:Explainable AI
User-centered Requirements	USR-031, USR-081, USR-085, USR-089
Technical Requirements	Req-AIMaaS-FUNC-1, Req-AIMaaS-FUNC-2, Req-AIMaaS-FUNC-5
Legal Requirements	LEG-TSD-092-103,
Socio-ethical Requirements	SOC-ETH-TSD-011, SOC-ETH-TSD-010, SOC-ETH-TSD-013, SOC-ETH-TSD-017, SOC-ETH-TSD-032; SOC-ETH-TSD-033, SOC-ETH-TSD-034
Piloting Summary	The dry run scenario of the XAI-By design module using open LiDAR datasets
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	ATHENA (Automotive dataset – demonstration on in-house simulator), PASEU (Automotive dataset – demonstration in real car)
Comments / Open issues	This scenario will demonstrate the use of XAI-by design techniques for training an AI model for the problem of LIDAR super-resolution in the automotive pilot use-case. The anticipated outcome is to demonstrate that no performance degradation is observed in relevant tasks (e.g., object detection, odometry) with low-cost, instead of high-cost, LIDAR sensors. At this stage, we are going to utilize open datasets. During the second stage the solution will be integrated into the CARLA ROS framework, while in the 3rd stage, it will be integrated into the real car.

UCS-006

Table 27: Use Case Scenario 006

Use Case Scenario Description	
Scenario ID	UCS-006
Title	Data providers dataset attributes creation
Scenario Description	This scenario will validate the flow between the issuer (certification authority) and holder (data provider) to create verifiable credentials with homomorphically encrypted attributes about the data provider dataset. This scenario includes the following consecutive sub-scenarios: • **UCS-006.01:** The holder generates FHE threshold keys (one public key and two private keys for HE threshold). • **UCS-006.02:** The issuer encrypts an attribute value with the public key from step 1. • **UCS-006.03:** The issuer uploads the encrypted attributes from step 2 to IPFS and an IPFS endpoint is obtained. • **UCS-006.04:** The issuer generates and issues an SSI credential including the IPFS endpoint from step 3 as an

	attribute. This SSI credential is received and stored by the holder. • **UCS-006.05:** The holder obtains the IPFS endpoint from the SSI credential received in step 4. • **UCS-006.06:** The holder downloads the encrypted attribute value from the IPFS endpoint from step 5. • **UCS-006.07:** The holder decrypts the encrypted attribute value from step 6 using the two private keys generated in step 1. As a result, the attribute value in clear is obtained.
Digital Solutions/Functionalities to be tested	SSI-HE: Issuer-Holder
Technical partners involved	TECNALIA
ABBs and/or inner modules of ABBs that are related	SSI-HE
User-centered Requirements	USR-020, USR-041, USR-046, USR-048, USR-026, USR-027, USR-107
Technical Requirements	Req-SSI-HE-FUNC-1, Req-SSI-HE-FUNC-2, Req-SSI-HE-FUNC-3, Req-SSI-HE-NFUNC-1
Legal Requirements	LEG-USR-001, LEG-USR-016, LEG-USR-018, LEG-USR-024, LEG-USR-026, LEG-USR-027, LEG-USR-028, LEG-USR-029, LEG-USR-054, LEG-USR-060, LEG-USR-071, LEG-USR-073, LEG-USR-095, LEG-TSD-003, LEG-TSD-005, LEG-TSD-011, LEG-TSD-012, Leg-TSD-013, LEG-TSD-014, LEG-TSD-019, LEG-TSD-020, LEG-TSD-021, LEG-TSD-031, LEG-TSD-033, LEG-TSD-042-050, LEG-TSD-058, LEG-TSD-066, LEG-TSD-071-075, LEG-TSD-077-078, LEG-TSD-109, LEG-TSD-110
Socio-ethical Requirements	SOC-ETH-TSD-008
Piloting Summary	The dry run scenario of the “offline” phase of the SSI-HE operation.
Main Persona(s) involved in the scenario	Data Provider
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, a PoC of the SSI-HE functionality will be demonstrated considering a standalone and local deployment. In the following pilot phases, the solution will be distributed and connected to other TRUSTEE ABBs. Pre-requisites: SSI schema and SSI credential definition are defined in advance (for the validation, a “generic” schema will be defined; it will be updated for the TRUSTEE use cases in the next phase).

UCS-007

Table 28: Use Case Scenario 007

Use Case Scenario Description	
Scenario ID	UCS-007
Title	The data provider's dataset attributes validation
Scenario Description	This scenario will validate the flow between the holder (data provider) and verifier (TRUSTEE Platform) to request verifiable proofs of dataset attributes, applying a specific feature searching operation over encrypted attributes (defined by a TRUSTEE Consumer) and finally securely decrypting the result of the search. This scenario includes the following consecutive sub-scenarios: • **UCS-007.01:** The verifier makes a proof request of the attribute to the holder. Automatically, the verifier receives a

	proof response from the holder with the requested attribute value. • **UCS-007.02:** The verifier obtains the IPFS endpoint from the proof response received in step 1. • **UCS-007.03:** The holder downloads the encrypted attribute value from the IPFS endpoint from step 2. • **UCS-007.04:** The verifier defines a specific input value which is encrypted with the public key (generated in step 1 in test UCS-006). • **UCS-007.05:** The verifier applies the searching functionality of the encrypted input from step 4 on the encrypted attribute value from step 3. The encrypted result of the search is obtained. • **UCS-007.06:** The verifier applies a partial decryption of the result from step 5 with one of the two private keys (generated in step 1 in test UCS-006). • **UCS-007.07:** The holder applies the additional partial decryption over the encrypted result from step 5 with the other private key (generated in step 1 in test UCS-006). • **UCS-007.08:** The verifier fusions the two partially decrypted results from steps 6 and 7 to obtain the decrypted result.
Digital Solutions/Functionalities to be tested	SSI-HE: Holder-Verifier
Technical partners involved	TECNALIA
ABBs and/or inner modules of ABBs that are related	SSI-HE
User-centered Requirements	USR-032, USR-020, USR-041, USR-046, USR-048, USR-026, USR-027, USR-107
Technical Requirements	Req-SSI-HE-FUNC-1, Req-SSI-HE-FUNC-4
Legal Requirements	LEG-USR-001, LEG-USR-016, LEG-USR-018, LEG-USR-024, LEG-USR-026, LEG-USR-027, LEG-USR-028, LEG-USR-029, LEG-USR-030, LEG-USR-054, LEG-USR-060, LEG-USR-071, LEG-USR-073, LEG-USR-095, LEG-TSD-003, LEG-TSD-005, LEG-TSD-011, LEG-TSD-012, Leg-TSD-013, LEG-TSD-014, LEG-TSD-019, LEG-TSD-020, LEG-TSD-021, LEG-TSD-031, LEG-TSD-033, EG-TSD-042-050, LEG-TSD-066, LEG-TSD-071-075, LEG-TSD-077-078, LEG-TSD-058, LEG-TSD-109, LEG-TSD-110
Socio-ethical Requirements	SOC-ETH-TSD-008
Piloting Summary	The dry run scenario of the "online" phase of the SSI-HE operation. The "offline" phase must have been executed in advance.
Main Persona(s) involved in the scenario	Data Provider
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, a PoC of the SSI-HE functionality will be demonstrated considering a standalone and local deployment. In the following pilot phases, the solution will be distributed and connected to other TRUSTEE ABBs. Pre-requisites: UCS-003 must be executed in advance.

UCS-008

Table 29: Use Case Scenario 008

Use Case Scenario Description	
Scenario ID	UCS-008

Title	Storing Transactions in permissioned blockchain
Scenario Description	This scenario will demonstrate recorded transactions/data within the permissioned blockchain. The entire blockchain network components will be demonstrated enabling to query all the recorded transactions or a specific set of transactions.
Digital Solutions/Functionalities to be tested	ATR:Consensus, ATR:MSP, ATR:Orderer, ATR:Peer, ATR:Ledger, ATR:Chaincode
Technical partners involved	ENU
ABBs and/or inner modules of ABBs that are related	ATR:PermissionedBlockchain
User-centered Requirements	USR-005, USR-045, USR-057, USR-004
Technical Requirements	Req-ATR-FUNC-4, Req-ATR-FUNC-1, Req-ATR-FUNC-2, Req-ATR-FUNC-3
Legal Requirements	LEG-TSD-003, LEG-TSD-005, LEG-TSD-005-006, LEG-TSD-012, LEG-TSD-013, LEG-TSD-019, LEG-TSD-023, LEG-TSD-035, LEG-TSD-042-047, LEG-TSD-066, LEG-TSD-072, LEG-TSD-087, LEG-TSD-088, LEG-TSD-089, LEG-USR-094-095, LEG-TSD-109, LEG-TSD-110, LEG-TSD-111-113, LEG-TSD-115
Socio-ethical Requirements	SOC-ETH-TSD-015, SOC-ETH-TSD-041
Piloting Summary	The dry run scenario of permissioned blockchain functionalities with recorded transactions in the ledger.
Main Persona(s) involved in the scenario	Data Provider, Model Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	The main aim of this scenario is to demonstrate key ATR functionalities and to determine what enhancements are required. Considering the suggestion and evaluation results from the 1st phase, an updated demonstration will be performed in the 2nd pilot phase.

UCS-009

Table 30: Use Case Scenario 009

Use Case Scenario Description	
Scenario ID	UCS-009
Title	Blockchain Monitoring
Scenario Description	This scenario will enable to monitor the inner functioning of blockchain.
Digital Solutions/Functionalities to be tested	ATR:Blockchain_Monitoring
Technical partners involved	ENU
ABBs and/or inner modules of ABBs that are related	ATR
User-centered Requirements	USR-057
Technical Requirements	Req-ATR-FUNC-5
Legal Requirements	LEG-TSD-005, LEG-TSD-035, LEG-TSD-042-043, LEG-TSD-046-048, LEG-TSD-045, LEG-TSD-109-110
Socio-ethical Requirements	SOC-ETH-TSD-15
Piloting Summary	The dry run scenario of monitoring the functioning of the ATR, such as transaction details, network performance, network activity, etc.

Main Persona(s) involved in the scenario	Data Provider, Model Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, the monitoring functionality, through the integration of an external tool, will be demonstrated. Other monitoring functionalities can be discussed based on the evaluation results and the enhancement can be performed in the 2nd phase.

UCS-010

Table 31: Use Case Scenario 010

Use Case Scenario Description	
Scenario ID	UCS-010
Title	Mouse automation capabilities of Robotic Process Automation
Scenario Description	This scenario will validate the capabilities of the RPA robot to simulate mouse events (mouse move, click etc.) and interpret the automation process file.
Digital Solutions/Functionalities to be tested	DA:RPAMouseActions
Technical partners involved	ADR
ABBs and/or inner modules of ABBs that are related	DA:RPA
User-centered Requirements	USR-048, USR-049, USR-051, USR-052, USR-055, USR-056, USR-115
Technical Requirements	Req-DA-FUNC-7, Req-DA-FUNC-9, Req-DA-FUNC-10, Req-DA-FUNC-8
Legal Requirements	LEG-TSD-042, LEG-TSD-007, LEG-TSD-025, LEG-TSD-026, LEG-TSD-027, LEG-TSD-030
Socio-ethical Requirements	SOC-ETH-TSD-008
Piloting Summary	Dry run scenario of RPA Robot performing mouse actions
Main Persona(s) involved in the scenario	Data Provider, Model Provider, Consumer
Foreseen Pilot Assistant	UCSC (Health dataset), ISS (Space dataset)
Comments / Open issues	For the 1st pilot phase, a PoC of the RPA robot will be presented. During the dry run of this scenario, the robot might not be able to detect the elements in the screen, which means that the mouse events might not be accurate. The functionality of the robot will continue to improve and an updated presentation will be performed in 2nd and 3rd Pilot Phase.

UCS-011

Table 32: Use Case Scenario 011

Use Case Scenario Description	
Scenario ID	UCS-011
Title	Distributed repository for content storing and sharing
Scenario Description	This scenario will demonstrate the key functionalities of the Knowledge Repository
Digital Solutions/Functionalities to be tested	KR:DistributedContentSharing
Technical partners involved	ENU
ABBs and/or inner modules of ABBs that are related	KR:PeerNodes

User-centered Requirements	USR-015, USR-084, USR-052, USR-030
Technical Requirements	Req-KR-FUNC-1
Legal Requirements	LEG-TSD-012-014, LEG-TSD-023-027, LEG-TSD-035, LEG-TSD-041-050, LEG-TSD-069-070, LEG-TSD-073, LEG-TSD-079, LEG-TSD-083
Socio-ethical Requirements	SOC-ETH-TSD-008, SOC-ETH-TSD-012; SOC-ETH-TSD-015, SOC-ETH-TSD-016, SOC-ETH-TSD-018, SOC-ETH-TSD-021, SOC-ETH-TSD-032
Piloting Summary	A cluster of 3 or more nodes will be formed in a private IPFS network. The nodes will communicate with each other for content storing and sharing. Additionally, key functionalities of the IPFS (KR) will be demonstrated, such as content adding, pinning contents, garbage collection feature, DHT, bootstrapping, etc.
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st pilot phase, the core functionalities of the IPFS will be demonstrated. Additional functionalities will be added in the 2nd Pilot phase based on the evaluation results and also partners contribution on two particular KR functionalities they are focusing on.

UCS-012

Table 33: Use Case Scenario 012

Use Case Scenario Description	
Scenario ID	UCS-012
Title	Monitoring Interface
Scenario Description	Emulation of the integration of KR (implemented utilising IPFS) with external monitoring tools/Dashboard
Digital Solutions/Functionalities to be tested	KR:Monitoring
Technical partners involved	ENU
ABBs and/or inner modules of ABBs that are related	KR:IPFSExporting
User-centered Requirements	USR-058
Technical Requirements	Req-KR-FUNC-2
Legal Requirements	LEG-TSD-001-002, LEG-TSD-006, LEG-TSD-011-014, LEG-TSD-020, LEG-TSD-023-027, LEG-TSD-032, LEG-TSD-035, LEG-TSD-043, LEG-TSD-079
Socio-ethical Requirements	SOC-ETH-TSD-015, SOC-ETH-TSD-025
Piloting Summary	In this scenario, a monitoring tool will be configured for integration with the knowledge repository. The tool should enable to obtain local node information, content details, content sharing information, peers, information, content search functionalities, content transfer history, and so on.
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	The outcome of the scenario will demonstrate the internal functioning of the Knowledge repository. Based on the evaluation results and suggestions from the 1st pilot phase, further functionalities will be added in the 2nd phase.

UCS-013

Table 34: Use Case Scenario 013

Use Case Scenario Description	
Scenario ID	UCS-013
Title	Create a user
Scenario Description	The user registers to TRUSTEE and the STM creates a profile of the registered user. The user can then be associated to agreements within the STM.
Digital Solutions/Functionalities to be tested	STM:CreateUser
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:InternalAPI, STM:UserManager, STM:GraphDB, STM:Ontology
User-centered Requirements	USR-001, USR-006, USR-029, USR-054, USR-110
Technical Requirements	Req-STM-FUNC-1, Req-STM-FUNC-8, Req-STM-NFUNC-2
Legal Requirements	LEG-TSD-001-062, LEG-TSD-064, LEG-TSD-071, LEG-TSD-085, LEG-TSD-087-091, LEG-TSD-107-113, LEG-USR-001-004, LEG-USR-007-016, LEG-USR-020-022, LEG-USR-027-033, LEG-USR-053, LEG-USR-057-059, LEG-USR-061, LEG-USR-063-066, LEG-USR-070, LEG-USR-072-075, LEG-USR-079-080, LEG-USR-094-099
Socio-ethical Requirements	SOC-ETH-TSD-002, SOC-ETH-TSD-005, SOC-ETH-USR-001, SOC-ETH-USR-002, SOC-ETH-USR-003, SOC-ETH-USR-004
Piloting Summary	Dry run of a user profile creation into TRUSTEE.
Main Persona(s) involved in the scenario	Data Provider, Model Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For M12, a PoC of the registration functionality will be demonstrated for the STM's internal use. The functionality might be replaced through the AM ABB and an updated demonstration of it could then be performed in the 2nd Pilot Phase based on the evaluation of the results of M12.

UCS-014

Table 35: Use Case Scenario 014

Use Case Scenario Description	
Scenario ID	UCS-014
Title	Authentication
Scenario Description	Should a user already exist within the STM's database, then the user should be identified by STM as authenticated/authorized to start creating agreements, signing agreements, or querying datasets.
Digital Solutions/Functionalities to be tested	STM:Authentication
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:InternalAPI, STM:UserManager, STM:GraphDB, STM:Ontology
User-centered Requirements	USR-001, USR-005, USR-006, USR-007, USR-029, USR-054, USR-057, USR-064, USR-110
Technical Requirements	Req-STM-FUNC-1, Req-STM-FUNC-7, Req-STM-FUNC-8, Req-STM-NFUNC-2
Legal Requirements	LEG-TSD-001-039, LEG-TSD-042-064, LEG-TSD-066, LEG-TSD-068-080, LEG-TSD-085, LEG-TSD-087-091LEG-TSD-108-113,

	LEG-USR-001-004, LEG-USR-007-016, LEG-USR-020-022, LEG-USR-027-033, LEG-USR-053, LEG-USR-057-059, LEG-USR-061, LEG-USR-063-066, LEG-USR-070, LEG-USR-072-075, LEG-USR-079-080, LEG-USR-094-099
Socio-ethical Requirements	SOC-ETH-TSD-002; SOC-ETH-USR-005; SOC-ETH-USR-008; SOC-ETH-USR-004; SOC-ETH-USR-001; SOC-ETH-USR-14
Piloting Summary	Dry run of a user logging and authenticating into TRUSTEE and using the STM.
Main Persona(s) involved in the scenario	Data Provider, Model Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For M12, a PoC of the authentication functionality will be demonstrated in the context of the STM's internal use without interacting with any other ABBs. This functionality might be replaced through the AM ABB and then integrated into the STM to create a user object within the STM. After M12, an updated demonstration of it could then be performed in the next phase based on the evaluation of the results of M12.

UCS-015

Table 36: Use Case Scenario 015

Use Case Scenario Description	
Scenario ID	UCS-015
Title	Creating a dataset agreement
Scenario Description	Before a Data Provider can share their dataset through the TRUSTEE Platform, the Data Provider must create an agreement that is linked to their dataset to define how the dataset should be processed. This agreement can then be signed by Data Consumers.
Digital Solutions/Functionalities to be tested	STM:DatasetAgreementCreation
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:AgreementManager, STM:InternalAPI, STM:GraphDB, STM:Ontology
User-centered Requirements	USR-001, USR-003, USR-005, USR-006, USR-007, USR-009, USR-010, USR-019, USR-029, USR-033, USR-040, USR-041, USR-042, USR-043, USR-054, USR-082, USR-100, USR-110
Technical Requirements	Req-STM-FUNC-1, Req-STM-FUNC-3, Req-STM-FUNC-7, Req-STM-FUNC-8, Req-STM-FUNC-9
Legal Requirements	LEG-TSD-001-0009, LEG-TSD-011-015, LEG-TSD-018-035, LEG-TSD-041-059, LEG-TSD-063-064, LEG-TSD-069, LEG-TSD-073-080, LEG-TSD-085-091, LEG-TSD-106-115, LEG-USR-001-004, LEG-USR-007-016, LEG-USR-020-022, LEG-USR-027-033, LEG-USR-053, LEG-USR-057-059, LEG-USR-061, LEG-USR-063-066, LEG-USR-070, LEG-USR-072-075, LEG-USR-079-080, LEG-USR-094-099
Socio-ethical Requirements	SOC-ETH-TSD-008, SOC-ETH-TSD-014, SOC-ETH-TSD-021, SOC-ETH-TSD-024, SOC-ETH-USR-035, SOC-ETH-049
Piloting Summary	Dry run of a Data Provider creating a dataset agreement for their to-be-shared dataset.
Main Persona(s) involved in the scenario	Data Provider
Foreseen Pilot Assistant	All

Comments / Open issues	For M12 a PoC is provided where it is assumed that the STM has access to a Data Provider's dataset metadata, such that the Data Provider can create an agreement for the dataset in an informed manner. This use case could be changed depending on the results of the integration phase with other ABBs.

UCS-016

Table 37: Use Case Scenario 016

Use Case Scenario Description	
Scenario ID	UCS-016
Title	Searching For Datasets to sign an Agreement for
Scenario Description	After registering their dataset(s) as part of the Distributed Data Sources of TRUSTEE, Data Providers can search over the datasets they have registered and select for which one they wish to sign an agreement
Digital Solutions/Functionalities to be tested	STM:DatasetSearch
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:DataManager, STM:UserManager, STM:InternalAPI
User-centered Requirements	USR-001, USR-003, USR-005, USR-019, USR-020, USR-029, USR-033, USR-036, USR-040, USR-041, USR-042, USR-045, USR-054, USR-057, USR-107, USR-110
Technical Requirements	Req-STM-NFUNC-1, Req-STM-NFUNC-2, Req-STM-FUNC-2
Legal Requirements	LEG-TSD-001-009, LEG-TSD-011-015, LEG-TSD-018-035, LEG-TSD-041-050, LEG-TSD-055-058, LEG-TSD-063-064, LEG-TSD-069, LEG-TSD-073-080, LEG-TSD-085-086, LEG-TSD-106-115, LEG-USR-001-002, LEG-USR-011, LEG-USR-014, LEG-USR-017, LEG-USR-019-022, LEG-USR-027-028, LEG-USR-030-033, LEG-USR-053, LEG-USR-060, LEG-USR-062; LEG-USR-071-072, LEG-USR-074, LEG-USR-094-099
Socio-ethical Requirements	SOC-ETH-USR-001, SOC-ETH-USR-005, SOC-ETH-USR-14, SOC-ETH-TSD-008, SOC-ETH-TSD-014, SOC-ETH-TSD-021, SOC-ETH-TSD-024, SOC-ETH-TSD-034, SOC-ETH-TSD-041
Piloting Summary	Dry run of a Data Provider searching through their registered datasets in order to select one for signing agreement for it
Main Persona(s) involved in the scenario	Data Provider
Foreseen Pilot Assistant	All
Comments / Open issues	For M12 a PoC is provided where it is assumed that the STM has access to a Data Provider's dataset metadata. For the PoC the STM performs its queries only within the STM ABB itself. The sequence of functions performed for this use case could change after integrating other ABBs.

UCS-017

Table 38: Use Case Scenario 017

Use Case Scenario Description	
Scenario ID	UCS-017

Title	View Information about a Dataset
Scenario Description	After querying a dataset through the SSI-HE and visualizing the query in the DASHBOARD the relevant dataset information is sent to the STM. This scenario refers to the emulation of a Consumer being able to see the selected dataset's metadata and the dataset's agreement in the STM, that must be signed should the Consumer want to process the dataset.
Digital Solutions/Functionalities to be tested	STM:ViewDataset
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:InternalAPI, STM:AgreementManager, STM:DataManager, STM:UserManager, STM:GraphDB, STM:Ontology
User-centered Requirements	USR-001, USR-005, USR-019, USR-020, USR-029, USR-033, USR-040, USR-041, USR-042, USR-045, USR-054, USR-110
Technical Requirements	Req-STM-FUNC-2, Req-STM-NFUNC-1, Req-STM-NFUNC-2
Legal Requirements	LEG-TSD-001-003, LEG-TSD-005, LEG-TSD-011-12, LEG-TSD-014, LEG-TSD-021-023, LEG-TSD-035, LEG-TSD-063-064, LEG-TSD-069, LEG-TSD-071-072, LEG-TSD-075, LEG-TSD-077-078, LEG-TSD-085, LEG-TSD-089, LEG-TSD-107,LEG-TSD-109, LEG-USR-001-004, LEG-USR-007, LEG-USR-009, LEG-USR-011, LEG-USR-014, LEG-USR-024, LEG-USR-027, LEG-USR-029, LEG-USR-053, LEG-USR-068, LEG-USR-072-073, LEG-USR-094-099
Socio-ethical Requirements	SOC-ETH-USR-002, SOC-ETH-USR-004, SOC-ETH-005, SOC-ETH-USR-014, SOC-ETH-USR-015
Piloting Summary	Dry run of a Data Consumer selecting a dataset after performing a search query. The Data Consumer may then view more details of the selected dataset in order to sign an agreement for it.
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For M12 a PoC is provided where it is assumed that the STM has access to a Data Provider's dataset metadata. The STM has example datasets stored internally for PoC purposes, as there is no integration with other ABBs at M12. The sequence of functions performed for this use case could change after integrating other ABBs.

UCS-018

Table 39: Use Case Scenario 018

Use Case Scenario Description	
Scenario ID	UCS-018
Title	View Dataset Agreements
Scenario Description	During the use of the STM, both Data Providers and Data Consumers sign several agreements. This use case represents the scenario where either user role would like to view all of their currently accepted/signed agreements.
Digital Solutions/Functionalities to be tested	STM:ViewDatasetAgreements
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:InternalAPI, STM:AgreementManager, STM:GraphDB, STM:Ontology
User-centered Requirements	USR-006, USR-041, USR-042, USR-043

Technical Requirements	Req-STM-FUNC-8, Req-STM-FUNC-10
Legal Requirements	LEG-TSD-001-003, LEG-TSD-005, LEG-TSD-011-12, LEG-TSD-014, LEG-TSD-021-023, LEG-TSD-035, LEG-TSD-063-064, LEG-TSD-069, LEG-TSD-071-072, LEG-TSD-075, LEG-TSD-077-078, LEG-TSD-085, LEG-TSD-089, LEG-TSD-107,LEG-TSD-109, LEG-USR-001-004, LEG-USR-007, LEG-USR-009, LEG-USR-011, LEG-USR-014, LEG-USR-024, LEG-USR-027, LEG-USR-029, LEG-USR-053, LEG-USR-068, LEG-USR-072-073, LEG-USR-094-099
Socio-ethical Requirements	SOC-ETH-USR-002, SOC-ETH-USR-004, SOC-ETH-005, SOC-ETH-USR-014, SOC-ETH-USR-015
Piloting Summary	Dry run of a Data Provider or Data Consumer viewing their agreements for overview and monitoring reasons.
Main Persona(s) involved in the scenario	Data Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For M12 a PoC is provided within which the STM stores the agreements of each user (Data Provider and Data Consumer). Currently, a listing of the agreements is displayed. This functionality is enhanced during the next phase to provide more in-depth information when monitoring agreements.

UCS-019

Table 40: Use Case Scenario 019

Use Case Scenario Description	
Scenario ID	UCS-019
Title	Accepting And Signing a Dataset Agreement
Scenario Description	After the Consumer has viewed the details of a dataset and the dataset's associated agreement, the Consumer may agree to the conditions within the dataset's agreement and sign the agreement to start with the data processing.
Digital Solutions/Functionalities to be tested	STM:SignDatasetAgreement
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:InternalAPI, STM:AgreementManager, STM:GraphDB, STM:Ontology
User-centered Requirements	USR-001, USR-003, USR-005, USR-019, USR-029, USR-033, USR-041, USR-042, USR-045, USR-054, USR-110
Technical Requirements	Req-STM-FUNC-1, Req-STM-FUNC-4, Req-STM-FUNC-8, Req-STM-FUNC-9
Legal Requirements	LEG-TSD-001-002, LEG-TSD-005, LEG-TSD-011-014, LEG-TSD-017, LEG-TSD-020-021, LEG-TSD-051-059, LEG-TSD-063-064, LEG-TSD-069, LEG-TSD-075, LEG-TSD-078-080, LEG-TSD-087-091, LEG-TSD-108-112, LEG-TSD-114-115, LEG-USR-001-004, LEG-USR-007-015, LEG-USR-027-033, LEG-USR-053, LEG-USR-057-061, LEG-USR-069, LEG-USR-072-073, LEG-USR-075, LEG-USR-079-080, LEG-USR-094-099
Socio-ethical Requirements	SOC-ETH-TSD-008, SOC-ETH-TSD-014, SOC-ETH-TSD-021, SOC-ETH-TSD-024, SOC-ETH-034, SOC-ETH-TSD-041, SOC-ETH-USR-002, SOC-ETH-USR-003
Piloting Summary	Dry run of a Consumer agreeing to the conditions of processing a dataset and therefore signing the agreement.

Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For M12 a PoC is provided where a basic signing is performed to demonstrate that two parties (Data Provider and Consumer) sign an agreement, such that both parties agree to the processing of a dataset. This use case will then be improved with trusted eSignatures and could be modified depending on the results of the integration with other ABBs.

UCS-020

Table 41: Use Case Scenario 020

Use Case Scenario Description	
Scenario ID	UCS-020
Title	View Profiles
Scenario Description	During the use of the STM, a user (Data Provider, and Consumer) can view their own profile and account information.
Digital Solutions/Functionalities to be tested	STM:ViewProfile
Technical partners involved	FUJITSU
ABBs and/or inner modules of ABBs that are related	STM:Frontend, STM:InternalAPI, STM:AgreementManager, STM:UserManager, STM:GraphDB, STM:Ontology
User-centered Requirements	USR-006, USR-019, USR-020, USR-040
Technical Requirements	Req-STM-FUNC-8, Req-STM-FUNC-9
Legal Requirements	LEG-TSD-001-039, LEG-TSD-049, LEG-TSD-042-050, LEG-TSD-054-059, LEG-TSD-063, LEG-TSD-066-103, LEG-TSD-106-110, LEG-TSD-111-113, LEG-USR-001-055, LEG-USR-057-099
Socio-ethical Requirements	SOC-ETH-USR-001; SOC-ETH-TSD-007; SOC-ETH-TSD-014; SOC-ETH-TSD-022; SOC-ETH-TSD-023; SOC-ETH-TSD-024; SOC-ETH-USR-003
Piloting Summary	Dry run of a Data Provider or Consumer viewing their account details
Main Persona(s) involved in the scenario	Data Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For M12 a PoC is provided within which the STM stores the agreements of each user (Data Provider and Consumer). Currently, a user can view their own profile.

UCS-021

Table 42: Use Case Scenario 021

Use Case Scenario Description	
Scenario ID	UCS-021
Title	AI trustwothiness Project Management
Scenario Description	This scenario will validate the correct project management of an AI trustworthiness assessment assigned task.
Digital Solutions/Functionalities to be tested	Project&PhaseManagementModule
Technical partners involved	RINA-C
ABBs and/or inner modules of ABBs that are related	TAI-SDF:OperationModule

User-centered Requirements	USR-037, USR-093, USR-091, USR-005, USR-057, USR-043
Technical Requirements	Req-TAI-SDF-NFUNC-1, Req-TAI-SDF-FUNC-3, Req-TAI-SDF-NFUNC-1, Req-TAI-SDF-FUNC-1
Legal Requirements	LEG-USR-081-093, LEG-TSD-092-103, LEG-USR-023
Socio-ethical Requirements	SOC-ETH-TSD-040; SOC-ETH-TSD-041
Piloting Summary	Dry run scenario of project management chain using Project&PhaseManagementModule
Main Persona(s) involved in the scenario	Model Provider, Consumer
Foreseen Pilot Assistant	ATHENA (Automotive dataset)
Comments / Open issues	-

UCS-022

Table 43: Use Case Scenario 022

Use Case Scenario Description	
Scenario ID	UCS-022
Title	Checklist compiling process
Scenario Description	This scenario will validate the correct checklist compiling process from the developer point of view. This part is the most important part since compilation of questionnaires allows the system to generate a score for the trustworthiness.
Digital Solutions/Functionalities to be tested	QuestionManagementModule ResponseManagementModule ComputedQuestionManagementModule
Technical partners involved	RINA-C
ABBs and/or inner modules of ABBs that are related	TAI-SDF:SurveyModule
User-centered Requirements	USR-037, USR-093, USR-091, USR-005, USR-043, USR-057
Technical Requirements	Req-TAI-SDF-FUNC-2, Req-TAI-SDF-NFUNC-1
Legal Requirements	LEG-TSD-093-103; LEG-USR-019-023;
Socio-ethical Requirements	SOC-ETH-TSD-040; SOC-ETH-TSD-041; SOC-ETH-TSD-041
Piloting Summary	Dry run scenario of User checklist compiling using SurveyModule
Main Persona(s) involved in the scenario	Model Provider, Consumer
Foreseen Pilot Assistant	ATHENA (Automotive dataset)
Comments / Open issues	For the 1st Pilot Phase, a PoC of the of the checklist compiling process will be demonstrated.

UCS-023

Table 44: Use Case Scenario 023

Use Case Scenario Description	
Scenario ID	UCS-023
Title	Questionnaire structure
Scenario Description	This scenario will validate the capability of the tool to generate a score for trustworthiness.
Digital Solutions/Functionalities to be tested	ReportModule
Technical partners involved	RINA-C
ABBs and/or inner modules of ABBs that are related	TAI-SDF: SurveyModule

User-centered Requirements	USR-037, USR-093, USR-091, USR-005, USR-043, USR-057
Technical Requirements	Req-TAI-SDF-FUNC-3, Req-TAI-SDF-NFUNC-1
Legal Requirements	LEG-TSD-092-103, LEG-USR-023, LEG-USR-081-093
Socio-ethical Requirements	SOC-ETH-TSD-040; SOC-ETH-TSD-041; SOC-ETH-TSD-041
Piloting Summary	The dry-run scenario of reporting using SurveyModule
Main Persona(s) involved in the scenario	Model Provider, Consumer
Foreseen Pilot Assistant	ATHENA (Automotive dataset)
Comments / Open issues	For the 1st Pilot Phase, a PoC of the of questionnaire structure and score evaluation will be demonstrated. The functionality will be further enhanced and demonstration of it will be performed in the 2nd Pilot Phase based on the evaluation of the results of the 1st Phase.

UCS-024

Table 45: Use Case Scenario 024

Use Case Scenario Description	
Scenario ID	UCS-024
Title	Authentication of TRUSTEE end-user
Scenario Description	This scenario is focusing on validating the processes of onboarding and registering users in the TRUSTEE ecosystem through the AM ABB. The scenario consists of the following sub-scenarios: • **UCS-024.01 – unregistered user login:** the login process is based on OIDC/OAuth2.0 protocol. New users are requested first onboard. • **UCS-024.02 – new user registration:** Proof of identity of users is being performed through bridging the eIDAS system and relevant attributes are issued as SSI VCs by the AM to the user to be stored at the user's wallet. • **UCS-024.03 – register user login:** the request is initiated by ABB Dashboard over OIDC/OAuth2.0 and the AM verifies the user' SSI ID.
Digital Solutions/Functionalities to be tested	AM:SSI Agent API, AM: Authentication Agent API, AM: Technology Collaboration, AM:eIDAS
Technical partners involved	INQBIT
ABBs and/or inner modules of ABBs that are related	ABB:AM (potential ABB:Dashboard, otherwise a mock-up)
User-centered Requirements	USR-003, USR-005, USR-008, USR-011, USR-019, USR-039, USR-044, USR-057, USR-054, USR-100
Technical Requirements	Req-AM-FUNC-1, Req-AM-FUNC-2, Req-AM-FUNC-3, Req-AM-FUNC-4, Req-AM-FUNC-5
Legal Requirements	LEG-TSD-001, LEG-TSD-002, LEG-TSD-003, LEG-TSD-005, LEG-TSD-006, LEG-TSD-012, LEG-TSD-013, LEG-TSD-014, LEG-TSD-019, LEG-TSD-020, LEG-TSD-021, LEG-TSD-022, LEG-TSD-035, LEG-TSD-036, LEG-TSD-041, LEG-TSD-042, LEG-TSD-043, LEG-TSD-044, LEG-TSD-045, LEG-TSD-046, LEG-TSD-047, LEG-TSD-048, LEG-TSD-049, LEG-TSD-050, LEG-TSD-051, LEG-TSD-052, LEG-TSD-053, LEG-TSD-054, LEG-TSD-055, LEG-TSD-056, LEG-TSD-057, LEG-TSD-058, LEG-TSD-059, LEG-TSD-060, LEG-TSD-061.
Socio-ethical Requirements	SOC-ETH-TSD-001, SOC-ETH-002, SOC-ETH-010, SOC-ETH-013, SOC-ETH-016, SOC-ETH-032, SOC-ETH-033, SOC-ETH-043.

Piloting Summary	Dry run scenario of User Authentication utilising OIDC and SSI over eIDAS
Main Persona(s) involved in the scenario	Data Provider, Model Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, a PoC of the registration functionality will be demonstrated. The functionality will be further enhanced and an updated demonstration of it will be performed in the 2nd Pilot Phase based on the evaluation of the results of the 1st Phase.

UCS-025

Table 46: Use Case Scenario 025

Use Case Scenario Description	
Scenario ID	UCS-025
Title	DPIA services
Scenario Description	This scenario will validate the processes involved during the provision of DPIA services to TRUSTEE users, including accessing the service and authenticate to the service, initiating new DPIA, working on existing DPIA, reviewing, approving or rejecting DPIA
Digital Solutions/Functionalities to be tested	DPIA:Front-end (Processes, Templates, Knowledge Base) DPIA:Back-end (RESTful API)
Technical partners involved	INQBIT
ABBs and/or inner modules of ABBs that are related	ABB:DPIA, ABB:AM, ABB:Dashboard
User-centered Requirements	USR-001, USR-003, USR-005, USR-006, USR-007, USR-009, USR-010, USR-029, USR-033, USR-036, USR-038, USR-040, USR-041, USR-042, USR-045, USR-046, USR-054, USR-057, USR-100, USR-107.
Technical Requirements	Req-DPIA-FUNC-1, Req-DPIA-FUNC-2, Req-DPIA-FUNC-3, Req-DPIA-FUNC-4
Legal Requirements	LEG-TSD-001, LEG-TSD-002, LEG-TSD-003, LEG-TSD-005, LEG-TSD-006, LEG-TSD-008, LEG-TSD-009, LEG-TSD-011, LEG-TSD-012, LEG-TSD-013, LEG-TSD-014, LEG-TSD-015, LEG-TSD-016, LEG-TSD-017, LEG-TSD-019, LEG-TSD-021, LEG-TSD-022, LEG-TSD-035, LEG-TSD-036, LEG-TSD-037, LEG-TSD-039, LEG-TSD-041, LEG-TSD-042, LEG-TSD-043, LEG-TSD-044, LEG-TSD-045, LEG-TSD-046, LEG-TSD-047, LEG-TSD-048, LEG-TSD-049, LEG-TSD-050, LEG-TSD-103, LEG-TSD-104, LEG-TSD-105, LEG-TSD-107, LEG-TSD-111.
Socio-ethical Requirements	SOC-ETH-TSD-001, SOC-ETH-002, SOC-ETH-010, SOC-ETH-013, SOC-ETH-016, SOC-ETH-032, SOC-ETH-033, SOC-ETH-043.
Piloting Summary	Dry run scenario of DPIA for authenticated TRUSTEE users
Main Persona(s) involved in the scenario	Data Provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, a PoC of the login functionality will be demonstrated. The functionality will be further enhanced and an updated demonstration of it will be performed in the 2nd Pilot Phase based on the evaluation of the results of the 1st Phase.

UCS-026

Table 47: Use Case Scenario 026

Use Case Scenario Description	
Scenario ID	UCS-026
Title	Support to Developers in the form of tutorials, instructions, and common practices relevant for the TRUSTEE Platform functioning and evolution
Scenario Description	Accessing the selected supporting materials offered by TRUSTEE OneSS by the Developer
Digital Solutions/Functionalities to be tested	OneSS:wiki
Technical partners involved	ENT + other project partners authoring the Dev handbook
ABBs and/or inner modules of ABBs that are related	OneSS:wiki
User-centered Requirements	USR-058, USR-063
Technical Requirements	Req-OneSS-NFUNC-StdAPIs Req-OneSS-NFUNC-StdAPIDoc
Legal Requirements	LEG-TSD-018, LEG-TSD-019, LEG-TSD-047, LEG-TSD-049, LEG-TSD-083, LEG-TSD-085, LEG-TSD-097, LEG-TSD-099,
Socio-ethical Requirements	SOC-ETH-TSD-022, SOC-ETH-TSD-024, SOC-ETH-TSD-031, SOC-ETH-TSD-038, SOC-ETH-TSD-042
Piloting Summary	When a new Developer starts the work the most relevant and critical information is to be found in the Developer Handbook hosted on https://github.com/Trustee-Horizon/Wiki
Main Persona(s) involved in the scenario	Developer
Foreseen Pilot Assistant	All
Comments / Open issues	Demonstrate the support and collection of the most critical information for the Developer including not only instructions but also summarizing personnel contacts on the main TRUSTEE Solution Building Blocks to facilitate the collaboration. In the 1st phase the initial wiki will be presented including Introduction, Onboarding, Development Process, TRUSTEE Architecture, Solution Building Blocks docs, Best Practices

UCS-027

Table 48: Use Case Scenario 027

Use Case Scenario Description	
Scenario ID	UCS-027
Title	HE keys generation
Scenario Description	This scenario will validate creation of the HE Keys needed for the confidential operation.
Digital Solutions/Functionalities to be tested	HEDF:keys generation
Technical partners involved	TECNALIA
ABBs and/or inner modules of ABBs that are related	HEDF
User-centered Requirements	USR-005
Technical Requirements	Req-HEDF-FUNCT-2
Legal Requirements	LEG-USR-066, LEG-TSD-003, LEG-TSD-005, LEG-TSD-012, LEG-TSD-014, LEG-TSD-019, LEG-TSD-021, LEG-TSD-031, LEG-TSD-033, LEG-043-047

Socio-ethical Requirements	SOC-ETH-TSD-008; SOC-ETH-TSD-009
Piloting Summary	Dry run scenario of the He keys generation in HEDF
Main Persona(s) involved in the scenario	Data provider, Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, the HE keys are generated without using the secure enclave. For the 1st Pilot Phase, everything related to HEDF HE layer happens on the same machine.

UCS-028

Table 49: Use Case Scenario 028

Use Case Scenario Description	
Scenario ID	UCS-028
Title	HE data encryption
Scenario Description	This scenario will validate the correct HE encryption of different datasets
Digital Solutions/Functionalities to be tested	HEDF:data encryption
Technical partners involved	TECNALIA
ABBs and/or inner modules of ABBs that are related	HEDF
User-centered Requirements	USR-003; USR-005; USR-009, USR-045, USR-102
Technical Requirements	Req-HEDF-FUNCT-2
Legal Requirements	LEG-TSD-003, LEG-TSD-005, LEG-TSD-012, LEG-TSD-014, LEG-TSD-019, LEG-TSD-021, LEG-TSD-031, LEG-TSD-033, LEG-TSD-043-047, LEG-TSD-069, LEG-TSD-074, LEG-USR-077, LEG-TSD-085, LEG-TSD-110, LEG-USR-001, LEG-USR-003, LEG-USR-024, LEG-USR-026, LEG-USR-066, LEG-USR-073-074, ,LEG-USR-079
Socio-ethical Requirements	SOC-ETH-TSD-008
Piloting Summary	Dry run scenario of the HE data encryption process
Main Persona(s) involved in the scenario	Data Provider
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, everything related to HEDF HE layer happens on the same machine.

UCS-029

Table 50: Use Case Scenario 029

Use Case Scenario Description	
Scenario ID	UCS-029
Title	Operation on the HE domain
Scenario Description	This scenario will validate the correct execution of different operations on the HE domain considering different data sources.
Digital Solutions/Functionalities to be tested	HEDF:data operation
Technical partners involved	TECNALIA
ABBs and/or inner modules of ABBs that are related	HEDF
User-centered Requirements	USR-016, USR-029; USR-033; USR-077; USR-114

Technical Requirements	Req-HEDF-FUNC-1, Req-HEDF-FUNCT-2, Req-HEDF-FUNCT-5, Req-DL-FUNCT-6
Legal Requirements	LEG-TSD-003, LEG-TSD-005, LEG-TSD-012, LEG-TSD-014, LEG-TSD-019, LEG-TSD-021, LEG-TSD-031, LEG-TSD-033, LEG-TSD-043-048, LEG-TSD-064, LEG-TSD-066, LEG-TSD-069, LEG-TSD-073-074, LEG-TSD-077, LEG-TSD-085, LEG-TSD-109, LEG-TSD-110, LEG-USR-001 (if the use case will use real personal data – LEG-USR-002-053 apply), LEG-USR-066, LEG-USR-073-074, LEG-USR-079
Socio-ethical Requirements	SOC-ETH-TSD-008
Piloting Summary	Dry run scenario of the operation on the HE domain
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, healthcare and educational required operations will be validated. These include: arithmetic operations, comparisons, searches, regressions, domains interrelation For the 1st Pilot Phase, everything related to HEDF HE layer happens on the same machine.

UCS-030

Table 51: Use Case Scenario 030

Use Case Scenario Description	
Scenario ID	UCS-030
Title	HE data decryption
Scenario Description	This scenario will validate the correct HE decryption of different datasets
Digital Solutions/Functionalities to be tested	HEDF:result decryption
Technical partners involved	TECNALIA
ABBs and/or inner modules of ABBs that are related	HEDF
User-centered Requirements	USR-045
Technical Requirements	Req-HEDF-FUNCT-2
Legal Requirements	LEG-TSD-003, LEG-TSD-005, LEG-TSD-012, LEG-TSD-014, LEG-TSD-019, LEG-TSD-021, LEG-USR-028, LEG-TSD-031, LEG-TSD-033, LEG-TSD-043-047, LEG-TSD-069, LEG-TSD-074, LEG-USR-077, LEG-TSD-085, LEG-TSD-110, LEG-USR-001, LEG-USR-003, LEG-USR-024, LEG-USR-026, LEG-USR-066, LEG-USR-073-074, LEG-USR-079
Socio-ethical Requirements	SOC-ETH-TSD-008
Piloting Summary	Dry run scenario of the result decryption process
Main Persona(s) involved in the scenario	Consumer
Foreseen Pilot Assistant	All
Comments / Open issues	For the 1st Pilot Phase, the He decryption process is executed without using the secure enclave. For the 1st Pilot Phase, everything related to HEDF HE layer happens on the same machine.

SUMMARY OF DRY RUN USE CASE SCENARIOS

The thirty (30) Dry Run Use Case Scenarios provided above refer to the dry run of the initial set of functionalities provided by the various ABBs of the TRUSTEE Platform, as it can be seen through the UCS tables referencing the ABBs and/or inner modules of ABBs that are related with each Use Case Scenario. Figure 7 below presents a summary of the Dry Run Use Case Scenarios for the 1st Pilot Phase showcasing the interconnection between Use Case Scenarios and respective ABBs.

Dry Run Use Case Scenarios	Related ABBs											
	HEDF	AIMaaS	TAI-SDF	ATR	KR	AM	DPIA	DA	SSI-HE	OneSS	DL	STM
UCS-001											X	
UCS-002											X	
UCS-003											X	
UCS-004		X										
UCS-005		X										
UCS-006									X			
UCS-007									X			
UCS-008				X								
UCS-009				X								
UCS-010								X				
UCS-011					X							
UCS-012					X							
UCS-013												X
UCS-014												X
UCS-015												X
UCS-016												X
UCS-017												X
UCS-018												X
UCS-019												X
UCS-020												X
UCS-021			X									
UCS-022			X									
UCS-023			X									
UCS-024						X						
UCS-025							X					
UCS-026										X		
UCS-027	X											
UCS-028	X											
UCS-029	X											
UCS-030	X											

Figure 7: UCS - ABBs Mapping

STRATEGY

This section gathers and presents the Use Case Scenarios to be tested within the context of TRUSTEE Pilot Use Cases or alongside the assistance and feedback by Pilot Use Case Leaders, during the 1st Pilot Phase as well as the methodology to followed for the test and demonstration.

All TRUSTEE Pilot Use cases have some common objectives regarding the 1st Pilot Phase. Since this Phase concerns a Dry Run of the various scenarios demonstrating initial functionalities of the ABBs of the TRUSTEE Platform, the common objectives include:

- Assisting in the conduction of the Dry Run Scenario phase and providing insights
- Understanding/familiarisation with the basic, initial set of functionalities of the solution offered by the TRUSTEE Platform and the various ABBs.

PILOT USE CASES

ENERGY DOMAIN

USE CASE SCENARIOS TO BE TESTED

The Use Case Scenarios that will be tested in the 1st Pilot Phase, in which the Energy Pilot Use Case Leader will assist and provide feedback, include the scenarios with the following IDs: UCS-001, UCS-

006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.

OBJECTIVES OF PHASE 1

The main objectives of the Energy Pilot Use Case in Phase 1 are to assist in the Dry Run testing of the ABB functionalities that are included in the Use Case Scenarios and familiarize with the developed solutions and technologies, as showcased in Table 52, below.

Table 52: Objectives of the Energy Pilot Use Case for Pilot Phase 1

Pilot Phase 1 – Dry Run Scenario: Reference Architecture Design & Preparation			
Objective #	**Objective**	**Pilot Domain**	**Use Case Scenario**
PP1.EN.Obj01	Assistance in the Dry Run testing of the ABB functionalities	Energy	UCS-001, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.EN.Obj02	Familiarization with the developed solutions and technologies	Energy	UCS-001, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030

METHODOLOGY

Dry Run demonstrations of the Use Case Scenarios and the involved ABB functionalities will take place. Where applicable, the integration platform offered by FORTH will be utilized for the demonstration and dry run pilot testing of the scenarios. For the above-mentioned scenarios, the Energy Pilot Use Case Leader will be present (in an online mode) and offer their insights and feedback while familiarizing with the developed solutions. The participation of the Energy Pilot Use Case Leader during pilot testing of the 1st Phase will also assist in the familiarization of ABB Leaders as well as of the TRUSTEE consortium as a whole with the energy dataset and the workflow to be adopted in the upcoming Pilot Phases inside the Energy Pilot Use Case.

HEALTH DOMAIN

USE CASE SCENARIOS TO BE TESTED

The Use Case Scenarios that will be tested in the 1st Pilot Phase, in which the Health Pilot Use Case Leader will assist and provide feedback, include the scenarios with the following IDs: UCS-001, UCS-002, UCS-003, UCS-006, UCS-007, UCS-008, UCS-009, UCS-010, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.

OBJECTIVES OF PHASE 1

The main objectives of the Health Pilot Use Case in Phase 1 are to assist in the Dry Run testing of the ABB functionalities that are included in the Use Case Scenarios and familiarize with the developed solutions and technologies.

Additionally, some more detailed objectives, related to the Use Case Scenarios to be tested, are considered as follows: verification that HE searches can be performed on numbers; verification that HE searches can be performed with acceptable performance; verification that tabular search can work on tables; verify that the implemented semi-transitive closure search can work on tables; evaluate how much the extracted semantic concepts differ compared to the schema type of information already known by the Data Provider.

The objectives of the Health Pilot Use Case for the 1st Pilot Phase are summarized in Table 53, below.

Table 53: Objectives of the Health Pilot Use Case for Pilot Phase 1

Pilot Phase 1 – Dry Run Scenario: Reference Architecture Design & Preparation			
Objective #	**Objective**	**Pilot Domain**	**Use Case Scenario**
PP1.HE.Obj01	Assistance in the Dry Run testing of the ABB functionalities	Health	UCS-001, UCS-002, UCS-003, UCS-006, UCS-007, UCS-008, UCS-009, UCS-010, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.HE.Obj02	Familiarization with the developed solutions and technologies	Health	UCS-001, UCS-002, UCS-003, UCS-006, UCS-007, UCS-008, UCS-009, UCS-010, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.HE.Obj03	Verification that HE searches can be performed on numbers	Health	UCS-029
PP1.HE.Obj04	Verification that HE searches can be performed with acceptable performance	Health	UCS-029
PP1.HE.Obj05	Verification that tabular search can work on tables	Health	UCS-029
PP1.HE.Obj06	Verification that the implemented semi-	Health	UCS-029

	transitive closure search can work on tables		
PP1.HE.Obj07	Evaluation of how much the extracted semantic concepts differ compared to the schema type of information already known by the Data Provider	Health	UCS-002, UCS-003

METHODOLOGY

Dry Run demonstrations of the Use Case Scenarios and the involved ABB functionalities will take place. Where applicable, the integration platform offered by FORTH will be utilized for the demonstration and dry run pilot testing of the scenarios. For the above-mentioned scenarios, the Health Pilot Use Case Leader will be present (in an online mode) and offer their insights and feedback while familiarizing with the developed solutions. The participation of the Health Pilot Use Case Leader during pilot testing of the 1st Phase will also assist in the familiarization of ABB Leaders as well as of the TRUSTEE consortium as a whole with the health dataset and the workflow to be adopted in the upcoming Pilot Phases inside the Health Pilot Use Case. In addition, since health and education datasets are the first to be considered for exploring secondary and multi-disciplinary data use through the TRUSTEE Platform, the insights offered by the Health Pilot Use Case Leader will pave the way towards extracting knowledge from the fusion of multi-disciplinary datasets.

EDUCATION DOMAIN

USE CASE SCENARIOS TO BE TESTED

The Use Case Scenarios that will be tested in the 1st Pilot Phase, in which the Education Pilot Use Case Leader will assist and provide feedback, include the scenarios with the following IDs: UCS-001, UCS-002, UCS-003, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.

OBJECTIVES OF PHASE 1

The main objectives of the Education Pilot Use Case in Phase 1 are to assist in the Dry Run testing of the ABB functionalities that are included in the Use Case Scenarios and familiarize with the developed solutions and technologies.

Additionally, some more detailed objectives, related to the Use Case Scenarios to be tested, are considered as follows: evaluation of HE as the preferred option to use, for which type of data and computation; evaluation of how much the extracted semantic concepts differ compared to the schema type of information already known by the Data Provider; verification of the capabilities of HE in computing basic mathematical operations with some of the provided data; evaluation of the size and capacity of computation once HE is used before such basic operations are computed; verification that HE searches can be performed on numbers; verification that such search performances are acceptable and the that a search result is obtained within an acceptable time frame; verification that search can be performed on large datasets.

The objectives of the Health Pilot Use Case for the 1st Pilot Phase are summarized in Table 54, below.

Table 54: Objectives of the Education Pilot Use Case for Pilot Phase 1

Pilot Phase 1 – Dry Run Scenario: Reference Architecture Design & Preparation			
Objective #	**Objective**	**Pilot Domain**	**Use Case Scenario**
PP1.ED.Obj01	Assistance in the Dry Run testing of the ABB functionalities	Education	UCS-001, UCS-002, UCS-003, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.ED.Obj02	Familiarization with the developed solutions and technologies	Education	UCS-001, UCS-002, UCS-003, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.ED.Obj03	Evaluation of HE as the preferred option to use, for which type of data and computation	Education	UCS-029
PP1.ED.Obj04	Evaluation of how much the extracted semantic concepts differ compared to the schema type of information already known by the Data Provider	Education	UCS-002, UCS-003
PP1.ED.Obj05	Verification of the capabilities of HE in computing basic mathematical operations with some of the provided data	Education	UCS-029
PP1.ED.Obj06	Evaluation of the size and capacity of computation once HE is used before such basic operations are computed	Education	UCS-028, UCS-029, UCS-030
PP1.ED.Obj07	Verification that HE searches can be performed on numbers	Education	UCS-029

PP1.ED.Obj08	Verification that such search performances are acceptable and the that a search result is obtained within an acceptable time frame	Education	UCS-028, UCS-029, UCS-030
PP1.ED.Obj09	Verification that search can be performed on large datasets	Education	UCS-029

METHODOLOGY

Dry Run demonstrations of the Use Case Scenarios and the involved ABB functionalities will take place. Where applicable, the integration platform offered by FORTH will be utilized for the demonstration and dry run pilot testing of the scenarios. For the above-mentioned scenarios, the Education Pilot Use Case Leader will be present (in an online mode) and offer their insights and feedback while familiarizing with the developed solutions. The participation of the Education Pilot Use Case Leader during pilot testing of the 1st Phase will also assist in the familiarization of ABB Leaders as well as of the TRUSTEE consortium as a whole with the education dataset and the workflow to be adopted in the upcoming Pilot Phases inside the Education Pilot Use Case. In addition, since health and education datasets are the first to be considered for exploring secondary and multi-disciplinary data use through the TRUSTEE Platform, the insights offered by the Education Pilot Use Case Leader will pave the way towards extracting knowledge from the fusion of multi-disciplinary datasets.

AUTOMOTIVE DOMAIN

USE CASE SCENARIOS TO BE TESTED

The Use Case Scenarios that will be tested in the 1st Pilot Phase, in which the Automotive Pilot Use Case Leader will assist and provide feedback, include the scenarios with the following IDs: UCS-001, UCS-002, UCS-004, UCS-005, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-021, UCS-022, UCS-023, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.

OBJECTIVES OF PHASE 1

The main objectives of the Automotive Pilot Use Case in Phase 1 are to assist in the Dry Run testing of the ABB functionalities that are included in the Use Case Scenarios and familiarize with the developed solutions and technologies.

More detailed objectives of Pilot Phase 1 in the Automotive Pilot Use Case are as follows:

- Assess the impact of incorporating HE in the parameter aggregation step of FL, on the performance of AI models trained via the FL paradigm.
- Demonstrate the benefit of utilizing the TRUSTEE Platform in a number of HE-enabled FL rounds, on the performance of AI models used in various automotive tasks.
- Demonstrate the benefits of XAI-by design / (federated) DU techniques for AI-based LIDAR super-resolution in the automotive pillar.

The objectives of the Automotive Pilot Use Case for the 1st Pilot Phase are summarized in Table 55, below.

Table 55: Objectives of the Automotive Pilot Use Case for Pilot Phase 1

Pilot Phase 1 – Dry Run Scenario: Reference Architecture Design & Preparation			
Objective #	**Objective**	**Pilot Domain**	**Use Case Scenario**
PP1.AU.Obj01	Assistance in the Dry Run testing of the ABB functionalities	Automotive	UCS-001, UCS-002, UCS-004, UCS-005, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-021, UCS-022, UCS-023, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.AU.Obj02	Familiarization with the developed solutions and technologies	Automotive	UCS-001, UCS-002, UCS-004, UCS-005, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-021, UCS-022, UCS-023, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.AU.Obj03	Assessment of the impact of incorporating HE in the parameter aggregation step of FL, on the performance of AI models trained via the FL paradigm	Automotive	UCS-004, UCS-028, UCS-029
PP1.AU.Obj04	Demonstration of the benefit of utilizing the TRUSTEE Platform in a number of HE-enabled FL rounds, on the performance of AI models used in various automotive tasks	Automotive	UCS-004
PP1.AU.Obj05	Demonstration of the benefits of XAI-by design / (federated) DU techniques for AI-based LIDAR super-resolution in the automotive pillar	Automotive	UCS-005

METHODOLOGY

Dry Run demonstrations of the Use Case Scenarios and the involved ABB functionalities will take place. Where applicable, the integration platform offered by FORTH will be utilized for the demonstration and dry run pilot testing of the scenarios.

At this stage, existing open automotive datasets and synthetic data will be utilized for model training and performance evaluation in both FL and XAI-by-design related scenarios. In UCS-004, the aim is to test the involved functionalities of HE-FL including setting up a group to collaborate via FL, tracking FL iterations and enabling the aggregation of local models, and different FL configurations (personalized, non-IID, constant vs variable aggregation weights).

In UCS-005, we will employ modern DU methodologies for designing and training interpretable and computationally and data efficient AI models for super-resolution of automotive LiDAR point-clouds. The benefits of improving raw data quality using DU models will be assessed on automotive LiDAR-based SLAM tasks, both in standalone and collaborative FL-based training scenarios.

For the rest of the above-mentioned scenarios, the Automotive Pilot Use Case Leader will be present (in an online mode) and offer their insights and feedback while familiarizing with the developed solutions. The participation of the Automotive Pilot Use Case Leader during pilot testing of the 1st Phase will also assist in the familiarization of ABB Leaders as well as of the TRUSTEE consortium as a whole with the automotive dataset and the workflow to be adopted in the upcoming Pilot Phases inside the Automotive Pilot Use Case.

SPACE DOMAIN

USE CASE SCENARIOS TO BE TESTED

The Use Case Scenarios that will be tested in the 1st Pilot Phase, in which the Space Pilot Use Case Leader will assist and provide feedback, include the scenarios with the following IDs: UCS-001, UCS-002, UCS-006, UCS-007, UCS-008, UCS-009, UCS-010, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.

OBJECTIVES OF PHASE 1

The main objectives of the Space Pilot Use Case in Phase 1 are to assist in the Dry Run testing of the ABB functionalities that are included in the Use Case Scenarios and familiarize with the developed solutions and technologies.

More detailed objectives of Pilot Phase 1 for the Space Pilot Use Case are as follows: validation of the homomorphic encryption by checking basic operations required by the Space Pilot Use Case, such as arithmetic operations and searches, can be performed in the encrypted domain; validation of the AM basic registration process; validation of the STM functionalities "Accepting/Signing the agreement".

The objectives of the Automotive Pilot Use Case for the 1st Pilot Phase are summarized in Table 56, below.

Table 56: Objectives of the Space Pilot Use Case for Pilot Phase 1

Pilot Phase 1 – Dry Run Scenario: Reference Architecture Design & Preparation			
Objective #	**Objective**	**Pilot Domain**	**Use Case Scenario**

PP1.SP.Obj01	Assistance in the Dry Run testing of the ABB functionalities	Space	UCS-001, UCS-002, UCS-006, UCS-007, UCS-008, UCS-009, UCS-010, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.SP.Obj02	Familiarization with the developed solutions and technologies	Space	UCS-001, UCS-002, UCS-006, UCS-007, UCS-008, UCS-009, UCS-010, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030
PP1.SP.Obj03	Validation of the homomorphic encryption by checking basic operations required by the Space Pilot Use Case, such as arithmetic operations and searches, can be performed in the encrypted domain	Space	UCS-029
PP1.SP.Obj04	Validation of the AM basic registration process	Space	UCS-024
PP1.SP.Obj05	Validation of the STM functionalities, such as "Accepting/Signing the agreement"	Space	UCS-016, UCS-017, UCS-018, UCS-019

METHODOLOGY

Dry Run demonstrations of the Use Case Scenarios and the involved ABB functionalities will take place. Where applicable, the integration platform offered by FORTH will be utilized for the demonstration and dry run pilot testing of the scenarios. For the above-mentioned scenarios, the Space Pilot Use Case Leader will be present (in an online mode) and offer their insights and feedback while familiarizing with the developed solutions. The participation of the Space Pilot Use Case Leader during pilot testing of the 1st Phase will also assist in the familiarization of ABB Leaders as well as of the TRUSTEE consortium as a whole with the space dataset and the workflow to be adopted in the upcoming Pilot Phases inside the Space Pilot Use Case.

TRUSTED MULTI-DISCIPLINARY DATA EXCHANGE

USE CASE SCENARIOS TO BE TESTED

The Use Case Scenarios that will be tested in the 1st Pilot Phase, in which the Trusted Multi-disciplinary Data Exchange Pilot Use Case Leader will assist and provide feedback, include the scenarios with the following IDs: UCS-001, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013,

UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.

OBJECTIVES OF PHASE 1

The main objectives of the Trusted Multi-disciplinary Data Exchange Pilot Use Case in Phase 1 are to assist in the Dry Run testing of the ABB functionalities that are included in the Use Case Scenarios and familiarize with the developed solutions and technologies, as showcased in Table 57, below.

Table 57: Objectives of the Trusted Multi-disciplinary Data Exchange Pilot Use Case for Pilot Phase 1

Pilot Phase 1 – Dry Run Scenario: Reference Architecture Design & Preparation			
Objective #	**Objective**	**Pilot Domain**	**Use Case Scenario**
PP1.MD.Obj01	Assistance in the Dry Run testing of the ABB functionalities	Multidisciplinary Data Exchange	UCS-001, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.
PP1.MD.Obj02	Familiarization with the developed solutions and technologies	Multidisciplinary Data Exchange	UCS-001, UCS-006, UCS-007, UCS-008, UCS-009, UCS-011, UCS-012, UCS-013, UCS-014, UCS-015, UCS-016, UCS-017, UCS-018, UCS-019, UCS-020, UCS-024, UCS-025, UCS-026, UCS-027, UCS-028, UCS-029, UCS-030.

METHODOLOGY

In the first Pilot Phase of the TRUSTEE Pan-European Pilot Campaign the focus of the Trusted Multi-disciplinary Data Exchange Pilot Use case is on discovering possible concrete cases on cross-pilot data exchange while in the next phases, more attention will be given to the sharing of data between the domains of the rest of the TRUSTEE Pilot Use cases, namely: Energy, Education, Health, Automotive, and Space. A first set of potential candidates for sharing data are the Education and Health pilots which will be further investigated by considering the outcomes of the 1st Pilot Phase and the insights and feedback provided by the rest of the Pilot Use Case Leaders during dry run pilot testing.

SUMMARY OF PILOT USE CASE OBJECTIVES

In Figure 8 below, a summary of the objectives of the Pilot Use Cases and their initial mapping to the respective Use Case Scenarios is provided. As it can be seen all objectives have been associated with one or more relevant Use Case Scenarios.

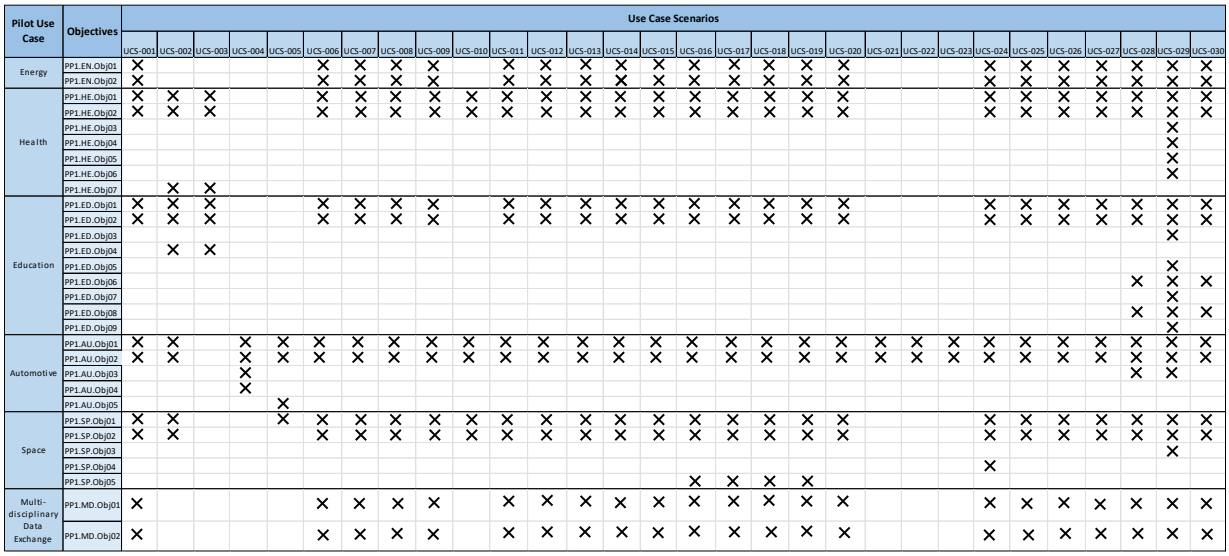

Figure 8: Pilot Use Case Objectives mapping to relevant Use Case Scenarios

KPIS

Several KPIs have been defined for the 1st Pilot Phase based on the Use Case Scenarios to be tested amd are presented in Table 58. In consideration of the Dry Run and PoC implementation of this phase, these KPIs are primarily centred on the implemented functionalities to be tested within the scenarios, and not the TRUSTEE Pilot Use Cases. In the upcoming Pilot Phases, when the Use Case Scenarios develop and become more complex, including intercommunication amongst ABBs, the Pilot Phase-specific KPIs will become relevant to the TRUSTEE Pilot Use Cases as well.

Table 58: KPIs of the 1st Pilot Phase

Pilot Phase 1 – Dry Run Scenario: Reference Architecture Design & Preparation					
KPI #	**KPI**	**ABB**	**Pilot Domain**	**Impact**	**Value**
PP1.1	Number of functionalities of TAI-SDF that support trustworthiness analysis	TAI-SDF	To be defined	High	0 to 25
PP1.2	Number of functionalities of TAI-SDF that simplifies trustworthiness analysis	TAI-SDF	To be defined	Medium	0 to 2
PP1.3	Mean time to conduct the trustworthiness analysis with Open Source SOTA	TAI-SDF	To be defined	Medium	Time
PP1.4	Number of users that execute tasks successfully	TAI-SDF	To be defined	High	The average number of users that execute tasks successfully
PP1.5	End-user satisfaction	TAI-SDF	To be defined	High	0 to 5

PP1.6	Number of operations/algorithms to be performed on the HE domain	HEDF	All	Very High	6
PP1.7	Number of different types of data to be considered for the confidential processing	HEDF	All	Very High	3
PP1.8	The suitable security level of the encryption process for data	HEDF	All	High	100%
PP1.9	Number of domains to be considered for the search functionality in the encrypted domain	HEDF	Health, Education	Very High	2
PP1.10	Number of different types of metadata to be considered for the datasets characterization	SSI-HE	All	Very High	3
PP1.11	Number of confidential attributes to be included in the same verifiable credential	SSI-HE	All	Medium	6
PP1.12	Suitable security level of the decryption process	SSI-HE	All	High	100%
PP1.13	The suitable security level of the encryption process for metadata	SSI-HE	All	High	100%
PP1.14	Impact of employing HE-enabled aggregation in FL rounds	AIMaaS: FL module	Automotive	High	5%
PP1.15	Complexity of AI models reduced via deep unrolling, without performance loss	AIMaaS: Explainable AI	Automotive	High	50%
PP1.16	Instances of performance enhancement due to the adoption of the FL training paradigm.	AIMaaS: FL module	Automotive	High	100%
PP1.17	Enabled access to summarized information on TRUSTEE architecture	OneSS (Handbook)	All	Medium	Wiki page about the stated topic
PP1.18	Enabled access to the description for each of the	OneSS (Handbook)	All	Medium	Wiki page for each solution building block

	TRUSTEE solution building blocks				
PP1.19	Enabled access to “best practices” related to development	OneSS (Handbook)	All	Medium	3 best practice topics
PP1.20	Enabled access to tutorials for developers	OneSS (Handbook)	All	Medium	3 tutorials
PP1.21	Total time in which an RPA Robot completes a specific action	DA	To be defined	Medium	1 to 2 seconds
PP1.22	Total errors of the RPA Robot when simulating mouse events	DA	To be defined	High	<2
PP1.23	Percentage of executed transactions without failures	ATR	To be defined	High	Percentage
PP1.24	The length of time for significant volume of recorded transactions	ATR	To be defined	High	Time
PP1.25	The length of time for block confirmation	ATR	To be defined	High	Time
PP1.26	Performance analysis for stability and inconsistencies	KR	To be defined	High	Time
PP1.27	GDPR compliance based on the agreement ontology	STM	All	High	70%
PP1.28	Frontend and backend: User identification and authentication	STM	All	High	100%
PP1.29	Frontend and backend: Creation of the dataset agreement	STM	All	High	100%
PP1.30	Frontend and backend: Signing the dataset agreement	STM	All	High	100%
PP1.31	Semantic dataset concepts stored in an RDF store	Data Lake	All	High	100%
PP1.32	Storing ability of public keys	Data Lake	All	Very High	100%

PP1.33	Total number of conducted DPIA(s)	DPIA	All	High	0 to 10
PP1.34	DPIAs Completed on Time	DPIA	All	High	50% to 75% (although in PP1 these numbers may be irrelevant)
PP1.35	DPIAs Requiring Mitigation	DPIA	All	Medium	<10% of DPIAs require mitigation
PP1.36	Average DPIA Review Time	DPIA	All	Medium	<30 minutes
PP1.37	DPIA compliance rate	DPIA	All	High	>50%
PP1.38	Privacy Risk Reduction	DPIA	All	High	>50% reduction in privacy risks
PP1.39	Authentication success rate	AM	All	Very High	95%-100%
PP1.40	Authentication Time	AM	All	Medium	~1 minute (for the purposes of PP1)
PP1.41	User Convenience	AM	All	Medium	High, Moderate, Low (simplified psychometric evaluation)
PP1.42	Security Effectiveness	AM	All	Very High	High – no security breach Moderate – Few security incidents with minimal impact
PP1.43	Authentication Error Rate	AM	All	High	<1% error rate
PP1.44	User Onboarding Time:	AM	All	Medium	<5 minutes (Indicative measure for the purposes and limitations of PP1)

INTEGRATION PLATFORM

An initial version of the cloud infrastructure to be used for dry run pilot testing during Pilot Phase 1 has been investigated by FORTH and is to be deployed between M13 and M18.

PARTICIPANTS

Considering the Dry Run Scenario and PoC implementation of Pilot Phase 1, the main actors are the technical partners of the TRUSTEE consortium developing and testing the dry run functionalities to be tested within the ABBS and the corresponding Use Case Scenarios. Nonetheless, since one of the

objectives of this phase is the familiarization of the partners leading the TRUSTEE Pilot Use Cases with the technologies and solutions developed within TRUSTEE, pilot leading partners are also involved in this phase as Foreseen Pilot Assistants with the main goal to observe and provide insights regarding the demonstration of functionalities of the dry run pilot testing and the integration platform used in Pilot Phase 1. As the project and the maturity of the ABBs progresses and the Use Case Scenarios become complex and Pilot Use Case specific in upcoming Pilot Phases, namely Pilot Phase 2, 3, and 4, the partners leading the TRUSTEE Pilot Use Cases will take on more active roles during pilot testing.

PARTNER ROLES AND RESPONSIBILITIES

This section presents the different roles and responsibilities assigned to TRUSTEE project Partners for the 1st Phase of the Pilot Campaign, as showcased in Figure 9 below.

The partners leading the Pilot Use Cases of TRUSTEE have been assigned the role of Pilot Use Case Rapporteurs apart from their role as Pilot Use Case Leaders, for the 1st Phase. Additionally, the partners leading the development of the ABBs, related to the Use Case Scenarios presented above, have been assigned the role of Technical Leaders of the respective Use Case Scenarios. The role of the Leader for the 1st Pilot Phase of the campaign has been assigned to HMU, which is also the Leader of WP5.

Pilot Phases	Role	PARTNER																						
		FORTH	HMU	ATHENA	PASEU	UoW	K3Y	ENU	ACCELI	UCSC	ENT	ADR	TECNALIA	RINA	ZORTENET	INOBIT	VTT	FUJITSU	EPL	HPE	HPE-CDS	ISS	CHECKWATT
Phase 1 - Dry Run Scenario: Reference Architecture Design & Preparation ALL Pilot Use Cases of TRUSTEE (ENERGY, HEALTH, EDUCATION, AUTOMOTIVE, SPACE, Trusted multi-disciplinary data exchange)	Pilot Use Case Leader				automotive					health							multi disciplinary		education			space	energy
	Pilot Phase Leader																						
	Pilot Use Case Rapporteur			automotive						health							multi disciplinary		education	automotive	automotive	space	energy
	Pilot Phase Rapporteur																						
	Technical Leader of Use Case Scenario			FL xAI				ATR KR	DL HEDF		OSS	DA	SSI-HE HEDF	TAI-SDF		AM DPIA		STM					
	Technical Assistant																						
	Legal Support																						

Figure 9: Partner Roles and Responsibilities for the 1st Pilot Phase

DATA ACQUISITION AND EXCHANGE

The data acquisition for Pilot Phase 1 will be conducted by the leaders of the respective Use Case Scenarios in cooperation with the assisting pilots (indicated in the section “Use Case Scenarios”).

The types of data, which are the basis for the input data from different use cases as well as the use cases themselves, have been initially described in D1.6 [4], and further in Part II of D2.1 [1].

The handling of the outcomes of the Use Case Scenarios foreseen for Pilot Phase 1 will be further defined and described in the form of a DHMP, as indicated in the Methodology for Pilot Campaign earlier in this document.

In the first step, the parties involved in Pilot Phase 1 will identify which questions from the DHMP are relevant and necessary to be addressed in this initial phase of the Pilot Campaign. The investigation is an ongoing task at the moment of submission of D5.1. A first version of the preliminary results of this investigation, regarding the relevance of the included questions to the use of Case Scenarios that will be tested during the 1st Pilot Phase, is presented in Annex however, as the investigation progress alongside

the run of Pilot Phase 1, these preliminary results are subject to further updates, enhancements, and modifications as the Pilot Phase will demand. Further, the ABB leaders in cooperation with Pilot Leaders will be requested to answer the questions to reflect the management of the data in Pilot Phase 1.

The outcomes of the analysis and collection of the information performed during the whole lifecycle of Pilot Phase 1 will be further presented in D5.2, and will further enrich the general Data Management Plan, namely D1.7, the next version of which will be submitted in M24.

LEGAL AND ETHICAL CONSIDERATIONS

The 1st Pilot Phase focuses on the "dry run" of selected functionalities of various ABBs, defined in D2.1, Part V "System Architecture" [1].

For each Use Case Scenario identified and described above as planned to be performed during Pilot Phase 1, a list of Legal and Socio-ethical requirements for TRUSTEE System Developers and TRUSTEE End-users in the context of creating and maintenance of the TRUSTEE solution, stemming from initial analysis and indication of these requirements in D2.1, has been presented. The aim of providing these requirements in the Use Case Scenarios is to ensure that further testing and validation of technical solutions in Pilot Phase 1 will take these requirements into consideration and appropriately address them, based on the current maturity of the ABB involved in the respective scenario. The responsible partner of each Use Case Scenario is in charge of fulfilling these requirements and is supported by the continuous assistance of UNIVIE and EPL.

The Dry Run is planned to involve the datasets provided by two TRUSTEE Pilot Use Cases: education as well as health, respectively led by EPL and UCSC. Both datasets, as reflected in D1.6 [4] and D2.1 [43], represent fictitious or dummy data, extracted and curated for the needs of the TRUSTEE project by the Pilot Use Case Leaders.

The data provided by the ABB Leaders represent technical, system data, such as source codes, dummy keys, and technical information pertaining to the components and their functionalities.

Based on the information obtained from the WP leader, at the moment of the start of Pilot Phase 1, there should be no real personal data included in the datasets provided by the selected Pilot Use Case Leaders for testing the functionalities as defined in Use Case Scenarios and involved in the performance of Pilot Phase 1. Any kind of changes in this context will be further investigated during Pilot Phase 1 and appropriately reflected in D5.2.

With regards to the Use Case Scenarios, which would generate the data from the behaviour of the users (members of the Pilot Leader's), such as, *inter alia*, UCS-010, the data obtained during the pilot phase will have to be processed and preserved in line with the requirements indicated in D1.6 [4] and, subsequently, in the section "Legal and Ethical Requirements" above. It is also relevant that the information that may be collected in Pilot Phase 1 by the means of Questionnaires indicated in "Quantitative Capturing Methods and Tools" as well as Observations, Depth Interviews, Storytelling and Brainstorming highlighted as part of "Qualitative Capturing Methods and Tools", which may also present personal information (i.e. personal data) of the pilot participants, is collected, curated and preserved in a data protection and data privacy compliant manner. The participants being natural persons should be informed and agree to the processing of their information, and the requirements stemming from the data protection laws with regard to the security of their data shall be fulfilled. The major requirements, which should be considered by the Partners conducting such qualitative and quantitative analysis in the scope

of the TRUSTEE Project, should therefore take into account the list of requirements indicated in D1.6 [4], and address them, where required. The requirements are as follows:

- Identification if the data are personal data.
- Implementation of the solutions addressing principles of personal data processing.
- Indication of a legal basis for processing such personal data.
- Provision of the information about data processing to the data subjects (pilot participants) – the most efficient manner of providing it should be considered.
- Established collaboration between controllers and processors of personal data, as well as recipients and third parties, e.g., in the form of a Joint Controllership Arrangement.
- Data Protection Impact Assessment conducted where required.
- Maintenance of the records of personal data processing activities by each of the Partners involved in this activity.
- Implementation of privacy and security by design approach.
- Applied security measures ensuring appropriate protection of the data of the pilot participants and other persons.
- Adherence to any local data protection laws, where identified by the Partners.
- Cross-border sharing of the data in compliance with the personal data protection laws

The detailed manner of addressing the requirements will be further investigated and defined during Pilot Phase 1 and reflected in further WP5 deliverables (D5.2-D5.5).

Additionally, for these types of data, the manner of handling and managing them should also be indicated for each Use Case Scenario in the DMHP and presented in D5.2 for Pilot Phase 1.

EXPECTED OUTCOMES OF PILOT PHASE 1

Following the guidelines set by the Definition Framework presented earlier in this document, this section discusses the foreseeable and expected outcomes of Pilot Phase 1 with regard to the maturity of the ABBs and with respect to the Dry Run scenario of this phase. The outcomes of the 1st Pilot Phase provide significant steps leading towards the successful realization of the next Pilot Phases, and ultimately the success of the entire TRUSTEE Pan-European Pilot Campaign.

ADOPTION AND USAGE

Pilot Phase 1 will provide initial knowledge transfer sessions and documentation that will introduce the TRUSTEE system components to end users, namely the TRUSTEE Pilot Use Case Leaders and, potentially, members of their organizations, and inform them about TRUSTEE's functionalities. Information about the system will be provided in an easy-to-access and straightforward way to facilitate the adoption of the system by the end users. Feedback will be gathered from the users to enhance the content in a user-centred way, supporting good user adoption. The system will also provide a first version of the DA demonstrating PoC and mock-up functionalities, such as a mock-up version of the Virtual Assistant (VA) module that will assist end users in navigation, and workflow automation and provide support in terms of frequently asked questions and wiki pages. The DA will be evaluated by the end users in terms of functionality suitability and easiness of use, and the end users' acceptance indicators will be provided. Another aspect aimed at increasing user adoption addressed by TRUSTEE is improving the trustworthiness and transparency of AI models. In Pilot Phase 1, a PoC implementation of the XAI-by-design functionality will be demonstrated via the Automotive pilot, showcasing the benefits of improving raw data in automotive LIDAR-based SLAM tasks using efficient and interpretable AI models. The work

done in Phase 1 of the Pilot Campaign will set a solid base for the further development of the XAI functionalities, which will be reported in the next pilot phases. All of the mentioned practices and their progress through the following phases of the Pilot Campaign will eventually increase TRUSTEE's user base.

DATA ECONOMY GROWTH

As discussed in the section "Expected Outcomes" related to the entire TRUSTEE Pilot Campaign, TRUSTEE will ultimately enable data economy growth by ensuring interoperability and better use of data by facilitating data sharing in a trustworthy and privacy-preserving way, hence also promoting collaboration between participants of the data economy. Although certain aspects of TRUSTEE's functionalities enabling such data sharing and collaboration, especially those related to privacy-preserving and data interoperability, will be demonstrated in their initial version during Pilot Phase 1, we have to observe data economy growth as one of the long-term outcomes. However, Pilot Phase 1 will pave the way towards its achievement, since validating the components of the TRUSTEE system and their initial functionalities will lead to the success of the later Pilot Phases and, eventually, bring successful support and added value for data economy participants.

SCALABILITY AND INTEGRATION

Within Pilot Phase 1, TRUSTEE's components (i.e., ABBs) will be validated in terms of their initial implemented set of PoC functionalities and will be assessed in the context of preparation for integration with other components and subsystems needed to realize the following Pilot Phases. The initial set of tests will be implemented and conducted to validate each component's functionalities included in the Dry Run scenarios.

SECURITY AND DATA PRIVACY

Pilot Phase 1 will enable the demonstration of selected aspects of the technologies enabling data privacy and security perseverance within TRUSTEE, such as certain scenarios using HE and FL, in accordance with the current implementation status of the relevant TRUSTEE components. In this Pilot Phase, PoC of TRUSTEE's security and privacy-enabling functionalities will be demonstrated in local and standalone deployment. Open datasets and synthetic data generated by partners' in-house environments will be used for demonstrating a PoC implementation of a privacy-preserving and secure HE-enabled FL training scheme showcasing, e.g., in the case of the Automotive Pilot Use Case, the benefits of employing FL for perception and visual odometry tasks, whereby only model parameters are shared while data are kept private, and parameter aggregation is based on secure HE-enabled computation. Validation of the basic operations in the encrypted domain needed by the pilots will be performed. Implemented functionalities of this Pilot Phase will pave the way for further progress and incremental incorporation of other solution blocks' functionalities in subsequent Pilot Phases.

TRAINING AND KNOWLEDGE TRANSFER

Knowledge transfer sessions will be organized and used for demonstrating the functionalities of the TRUSTEE components ready within Pilot Phase 1 to the end users. In addition to the knowledge transfer session, documentation will be provided, containing information about the TRUSTEE system, describing

its components, as well as the best practices, tutorials, and guidelines for TRUSTEE stakeholders, thus enabling straightforward access to the relevant information for the end users, especially developers.

CONCLUSION

This document presented the initial an overview of the TRUSTEE Pan-European Pilot Campaign including the initial Pilot Campaign Plan while it also provided the initial Definition Framework and the Evaluation Framework that describe the methodology to be followed for defining and evaluating the four (4) Pilot Phases and which will be used throughout the entire campaign. Additionally, in this document, the 1st Pilot Phase: Dry Run Scenario, which will run between M13 and M18, has been initially defined using the Definition Framework that was developed, whereas its results and outputs will be evaluated after its completion and will be reported in D5.2, due to be submitted on M19.

The methodologies described in this document for the definition and evaluation of the Pilot Phases set the groundwork and an initial set of guidelines to be followed and may be revised, enhanced, and complemented depending on additional demands that may arise during the project and the campaign itself.

REFERENCES

[1] TRUSTEE Consortium, "D2.1 Live doc conceptualization, use cases, and system architecture V1," 2023.

[2] T. Mäkelä and T. Leinonen, "Design framework and principles for learning environment co-design: Synthesis from literature and three empirical studies," *Buildings*, vol. 11, no. 12, 2021, doi: 10.3390/buildings11120581.

[3] A. M. Turner, B. Reeder, and J. Ramey, "Scenarios, personas and user stories: User-centered evidence-based design representations of communicable disease investigations," *J. Biomed. Inform.*, vol. 46, no. 4, pp. 575–584, 2013, doi: 10.1016/j.jbi.2013.04.006.

[4] TRUSTEE Consortium, "D1.6 TRUSTEE Data Management Plan available," 2023.

[5] E. Horvitz and D. Mulligan, "Data, privacy, and the greater good," *Science (80-.).*, vol. 349, no. 6245, pp. 253–255, 2015.

[6] N. B. Truong, K. Sun, G. M. Lee, and Y. Guo, "Gdpr-compliant personal data management: A blockchain-based solution," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, pp. 1746–1761, 2019.

[7] B. McMahan and D. Ramage, "Federated learning: Collaborative machine learning without centralized training data," *Google Res. Blog*, vol. 3, 2017.

[8] H. Pettenpohl, M. Spiekermann, and J. R. Both, "International data spaces in a nutshell," *Des. Data Spaces; Springer Cham, Switz.*, pp. 29–40, 2022.

[9] E. Curry, S. Scerri, and T. Tuikka, *Data Spaces*. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-030-98636-0.

[10] M. Jarke and C. Quix, "Federated Data Integration in Data Spaces," in *Designing Data Spaces*, Cham: Springer International Publishing, 2022, pp. 181–194. doi: 10.1007/978-3-030-93975-5_11.

[11] M. Farrell, "How to Conduct a Pilot Study," pp. 1–7, 2009, [Online]. Available: https://fuelcycle.com/blog/how-to-pilot-test/

[12] VICINITY Consortium, "D8.1 Business Scenarios & Evaluation Framework," no. 1, pp. 1–60, 2020.

[13] M. B. Dias, D. Dausey, E. Casman, and M. B. Dias, "The PREval (Pilot Research Evaluation) Framework," *Inf. Technol. Int. Dev.*, p. 1-4, 2011.

[14] A. Feidenheimer, "Use Cases and Pilots: Definition of Methodology Deliverable information Deliverable number and name D6.1 Use Cases and Pilots: Definition of Methodology Due date Month 15 Delivery Month 15 Work Package WP6 Lead Partner for deliverable PPMI Author Vilius S," 2019, [Online]. Available: http://europa.eu.int/

[15] A. Pearson, "Key performance indicators for nursing," *Int. J. Nurs. Pract.*, vol. 9, no. 6, pp. 337–337, 2003, doi: 10.1046/j.1440-172X.2003.00450.x.

[16] ITU-T, "ITU-T Recommendations under Study Group 20 responsibility." https://www.itu.int/ITU-T/recommendations/index_sg.aspx?sg=20 (accessed Jul. 03, 2023).

[17] R. PANNEERSELVAM, *Research methodology*. PHI Learning Pvt. Ltd., 2014.

[18] Voxco, "The Basics of Research Methodology: A Key to Quality Research," 2021. https://www.voxco.com/blog/what-is-research-methodology/ (accessed Mar. 20, 2007).

[19] S. Mehta, "Types of Research Methodology," 2023. https://eduvoice.in/types-research-methodology/ (accessed Mar. 20, 2007).

[20] Digital.gov, "Usability Starter Kit," *2022*. https://digital.gov/resources/digitalgov-user-experience-resources/digitalgov-user-experience-program-usability-starter-kit/?dg

[21] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models," *Manage. Sci.*, vol. 35, no. 8, pp. 982–1003, 1989.

[22] V. Venkatesh and H. Bala, "Technology acceptance model 3 and a research agenda on interventions," *Decis. Sci.*, vol. 39, no. 2, pp. 273–315, 2008.

[23] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Q.*, pp. 425–478, 2003.

[24] V. Venkatesh, J. Y. L. Thong, and X. Xu, "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology," *MIS Q.*, pp. 157–178, 2012.

[25] P. Kline, *Handbook of psychological testing*. Routledge, 2013.

[26] N. Chouliaras, G. Kittes, I. Kantzavelou, L. Maglaras, G. Pantziou, and M. A. Ferrag, "Cyber ranges and testbeds for education, training, and research," *Appl. Sci.*, vol. 11, no. 4, p. 1809, 2021.

[27] F. D. Davis and V. Venkatesh, "Toward preprototype user acceptance testing of new information systems: implications for software project management," *IEEE Trans. Eng. Manag.*, vol. 51, no. 1, pp. 31–46, 2004.

[28] T. Franke, C. Attig, and D. Wessel, "A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale," *Int. J. Human–Computer Interact.*, vol. 35, no. 6, pp. 456–467, 2019.

[29] J. Brook, "SUS: a" quick and dirty" usability scale," *Usability Eval. Ind.*, 1996.

[30] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Advances in psychology*, Elsevier, 1988, pp. 139–183.

[31] J. R. Lewis, "Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ," *ACM Sigchi Bull.*, vol. 23, no. 1, pp. 78–81, 1991.

[32] J. W. Creswell, *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson Education, Inc, 2012.

[33] S. M. S. Kabir, "Basic guidelines for research: An introductory approach for all disciplines," *B. Zo. Publ.*, vol. 4, no. 2, pp. 168–180, 2016.

[34] S. M. S. Kabir, "Methods OF Data Collection: An Introductory Approach for all Disciplines." Dhaka: Book Zone, 2016.

[35] H. Taherdoost, "Data Collection Methods and Tools for Research; A Step-by-Step Guide to Choose Data Collection Technique for Academic and Business Research Projects," *Int. J. Acad. Res. Manag.*, vol. 10, no. 1, pp. 10–38, 2021.

[36] P. Pandey and M. M. Pandey, *Research methodology tools and techniques*. Bridge Center, 2021.

[37] C. C. Wang and S. K. Geale, "The power of story: Narrative inquiry as a methodology in nursing research," *Int. J. Nurs. Sci.*, vol. 2, no. 2, pp. 195–198, 2015.

[38] D. J. Clandinin and F. M. Connelly, "Clandinin, D. Jean, and F. Michael Connelly, Narrative

Inquiry: Experience and Story in Qualitative Research. San Francisco: Jossey-Bass, 2000.," 2000.

[39] G. E. Mills, *Action research: A guide for the teacher researcher.* ERIC, 2000.

[40] J. W. Creswell and C. N. Poth, *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications, 2016.

[41] G. A. Bowen, "Document analysis as a qualitative research method," *Qual. Res. J.*, vol. 9, no. 2, pp. 27–40, 2009.

[42] Save the Children, "Monitoring, evaluation, accountability and learning (MEAL): 6 Methods of data collection and analysis. Research, reports and studies." https://resourcecentre.savethechildren.net/pdf/6_methods_of_data_collection.pdf/

[43] TRUSTEE Consortium, "D2.1 Live doc conceptualization, use cases and system architecture V1," 2023.

ANNEXES

ANNEX 1: DATA MANAGEMENT HANDLING PLANS (DMHP)

DMHP Questions			Use Case Scenario									
			UCS-001	UCS-002	UCS-003	UCS-004	UCS-005	UCS-006	UCS-007	UCS-008	UCS-009	UCS-010
DATA PRODUCTION AND STORAGE	*Types or categories of data generated/collected*	What types of research data are collected or generated during a specific pilot phase? Who (or which entity) will be responsible for deciding what data is collected or generated?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant
	Personal or non-personal data	Will the input or output data represent personal or non-personal data? What type of non-personal data will be collected at the pilot site in a specific pilot phase? What type of personal data will be collected?	Not relevant	To be determined	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant
	Dummy/fake or real data	Will the data be dummy/fake/fictitious/synthetic or real?	Not relevant	Relevant	To be determined	To be determined	To be determined	To be determined	To be determined	Relevant	Not relevant	Relevant
	Formats of the data	In which format will the data be collected (e.g., CSV, JSON, xls, PDF, ….)?	Not relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant
	Reproducibility of data	Please provide the information for validation and reuse of data and indicate if the data are foreseen as open access	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Not relevant	Not relevant
	Data size	Please provide the information about the estimated size of data provided as input, as well as foreseen size of the data produced.	Not relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	To be determined

	Software tools for creating/processing /visualising data	Which application/ABB will be tested in your pilot in this specific phase? What aspects/functionalities of the applications will be tested at your specific pilot location in a specific phase? Besides the ABBs indicated as to be tested in your pilot phase, what other software tools will be used for creating/processing/visualising data?	GENERAL COMMENT: It is relevant, nevertheless, this question seems to be answered by the UCS themselves, as each of them specifically indicates the ABB and functionality which will be tested.									
	Use of pre-existing data	Will you use pre-existing data? Yes / No / Uncertain. If so, please indicate what pre-existing data will be used.	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	To be determined
	Data storage and backup strategies	Please indicate what storage and backup strategies will be adopted for input and output data.	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	To be determined
	Purpose of data collection	Considering each type of data collected in the pilot phase, what is the purpose of their collection?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not Relevant	Not relevant	Relevant
ORGANISATION, DOCUMENTATION AND	*Standards for documentation of metadata*	What standards will be used for documentation and metadata (e.g., Digital Object Identifiers)? Is there a community standard for metadata sharing/integration?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Best practice/guidelines adopted for data management	Are there any best practices or guidelines which are foreseen to be applied in the context of organisation and documentation of the data and metadata in the pilot phases?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Tools for formatting data	What type of tools will you use to format data in the pilot phase?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	To be determined

	Directory and file naming convention used.	What directory and file naming convention will be used? Will you provide clear version numbers?	Not relevant	Relevant	Not relevant	Not relevant	Not relevant	Relevant	Not relevant	Not relevant	Not relevant	To be determined
DATA ACCESS	*Risks to data*	What main risks to data collected / produced during the pilot phase do you foresee? • Loss or destruction of data • Data breach • Loss of availability • Loss of integrity • Loss of confidentiality • Unauthorised alteration transmission and storage of data. Please provide any other major risks to data collected/produced at pilot sites	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Risk management	Have you prepared a formal risk assessment addressing each of the major risks to data security and potential solutions? If so, please share further information. If no/uncertain, please explain why.	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant
	Data access & requirements for access	Are there any concerns regarding access to your data? Yes / No	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Not relevant	Relevant
	Correct execution of the data access process	Please indicate a proper process which someone would need to take to access data collected/generated at pilot site during the pilot phase, as well as who is responsible for checking the correct execution of the access process. If data is confidential (e.g. personal data not already in the public domain, confidential business information or trade secrets), are there any appropriate security measures in place or any formal	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Not relevant	Relevant

		standards that you have to comply with?										
	Procedures to follow in the event of a data breach	Are there any specific data breach procedures which you foresee should be followed in the case of such an event?	Relevant	Relevant	Not relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Not relevant	Not relevant
DATA SHARING AND REUSE OF DATA	*Organization/labelling of Data for easy identification*	How will you organise or label the data to ensure that researchers may easily isolate fields of interest in their study?	Not relevant	Relevant	Relevant	Not relevant	Not relevant	Relevant	Not relevant	Not relevant	Not relevant	Not relevant
	Data Sharing & Audience for Data Sharing	Who can access data produced in the pilot and in a specific pilot phase?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Data Sharing Requirements	Are there any data sharing requirements which should be followed in the context of sharing the data produced/generated in the pilot in its specific pilot phase?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Not relevant	Relevant
	Re-use of data	Will the data produced or generated in the pilot during its specific phase made re-usable or openly accessible? Will the data be reproducible (i.e. able to be copied)?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	To be determined
	Audience for re-use	Who will use the data during the pilot? Who will use it afterwards?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant

	Restrictions on the re-use of data	Are there any restrictions regarding the entities that can re-use the data and for what purposes the data can be used?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Not relevant
	Publication of data	Do you plan to publish the data generated / collected in the pilot within its specific phase and if so, then where will you publish them?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	To be determined
DATA PRESERVATION AND ARCHIVING	*Archiving of data for preservation and long-term access*	How will the data produced within the pilot in its specific pilot phase be preserved for long-term access?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Time period for data retention	How long the data should or could be retained?	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	File formats of retained data	Please provide in what formats the data will be retained.	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Data archives	What type of data archives will be used to retain pilot generated/collected data?	Not relevant	Relevant	To be determined	To be determined	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Long-term maintenance of data (systems and procedures)	Please provide the details on envisioned systems and procedures for long-term maintenance of data.	Not relevant	To be determined	Not relevant	To be determined	To be determined	To be determined	To be determined	To be determined	To be determined	Relevant

(continues including UCS-011 – UCS-020)

DMHP Questions			**Use Case Scenario**									
			UCS-011	**UCS-012**	**UCS-013**	**UCS-014**	**UCS-015**	**UCS-016**	**UCS-017**	**UCS-018**	**UCS-019**	**UCS-020**
DATA PRODUCTION	*Types or categories of data generated/ collected*	What types of research data are collected or generated during a specific pilot phase? Who (or which entity) will be responsible for deciding what data is collected or generated?	To be determined	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant

	Personal or non-personal data	Will the input or output data represent personal or non-personal data? What type of non-personal data will be collected at the pilot site in a specific pilot phase? What type of personal data will be collected?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant
	Dummy/fake or real data	Will the data be dummy/fake/fictitious/synthetic or real?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant
	Formats of the data	In which format will the data be collected (e.g., CSV, JSON, xls, PDF, ….)?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant
	Reproducibility of data	Please provide the information for validation and reuse of data and indicate if the data are foreseen as open access	Relevant	Relevant	Not relevant	Relevant	Not relevant	Not relevant	Relevant		Relevant	Relevant
	Data size	Please provide the information about the estimated size of data provided as input, as well as foreseen size of the data produced.	Relevant	Not relevant	Not relevant	Not relevant	Not relevant	Relevant	Relevant		Relevant	Relevant
	Software tools for creating/processing /visualising data	Which application/ABB will be tested in your pilot in this specific phase? What aspects/functionalities of the applications will be tested at your specific pilot location in a specific phase? Besides the ABBs indicated as to be tested in your pilot phase, what other software tools will be used for creating/processing/visualising data?	GENERAL COMMENT: It is relevant, nevertheless, this question seems to be answered by the UCS themselves, as each of them specifically indicates the ABB and functionality which will be tested.									
	Use of pre-existing data	Will you use pre-existing data? Yes / No / Uncertain. If so, please indicate what pre-existing data will be used.	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant		Relevant	Relevant

	Data storage and backup strategies	Please indicate what storage and backup strategies will be adopted for input and output data.	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant		Relevant	Relevant
	Purpose of data collection	Considering each type of data collected in the pilot phase, what is the purpose of their collection?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant
ORGANISATION, DOCUMENTATION AND METADATA	*Standards for documentation of metadata*	What standards will be used for documentation and metadata (e.g., Digital Object Identifiers)? Is there a community standard for metadata sharing/integration?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant
	Best practice/guidelines adopted for data management	Are there any best practices or guidelines which are foreseen to be applied in the context of organisation and documentation of the data and metadata in the pilot phases?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant
	Tools for formatting data	What type of tools will you use to format data in the pilot phase?	Relevant	Relevant	Relevant	Not relevant	Not relevant	Not relevant	Relevant		Relevant	Relevant
	Directory and file naming convention used.	What directory and file naming convention will be used? Will you provide clear version numbers?	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant		Not relevant	Not relevant
DATA ACCESS	*Risks to data*	What main risks to data collected / produced during the pilot phase do you foresee? • Loss or destruction of data • Data breach • Loss of availability • Loss of integrity • Loss of confidentiality • Unauthorised alteration transmission and storage of data. Please provide any other major risks to data collected/produced at pilot sites	Not relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant		Relevant	Relevant

	Risk management	Have you prepared a formal risk assessment addressing each of the major risks to data security and potential solutions? If so, please share further information. If no/uncertain, please explain why.	Not relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant		Relevant	Relevant
	Data access & requirements for access	Are there any concerns regarding access to your data? Yes / No	Not relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant		Relevant	Relevant
	Correct execution of the data access process	Please indicate a proper process which someone would need to take to access data collected/generated at pilot site during the pilot phase, as well as who is responsible for checking the correct execution of the access process. If data is confidential (e.g. personal data not already in the public domain, confidential business information or trade secrets), are there any appropriate security measures in place or any formal standards that you have to comply with?	Not relevant	To be determined	To be determined	To be determined	To be determined	Not relevant	Relevant		Relevant	Relevant
	Procedures to follow in the event of a data breach	Are there any specific data breach procedures which you foresee should be followed in the case of such an event?	Not relevant	Not relevant	Not relevant	Relevant	Not relevant	Not relevant	To be determined		To be determined	To be determined
DATA SHARING AND REUSE OF DATA	*Organization/labelling of Data for easy identification*	How will you organise or label the data to ensure that researchers may easily isolate fields of interest in their study?	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant		Not relevant	Not relevant
	Data Sharing & Audience for Data Sharing	Who can access data produced in the pilot and in a specific pilot phase?	Relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant		Relevant	Relevant
	Data Sharing Requirements	Are there any data sharing requirements which should be followed in the context of sharing the data produced/generated in the pilot in its specific pilot phase?	Not relevant	Not relevant	Relevant	Relevant	Relevant	Not relevant	Not relevant		Not relevant	Relevant

	Re-use of data	Will the data produced or generated in the pilot during its specific phase made re-usable or openly accessible? Will the data be reproducible (i.e. able to be copied)?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Not relevant
	Audience for re-use	Who will use the data during the pilot? Who will use it afterwards?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Relevant
	Restrictions on the re-use of data	Are there any restrictions regarding the entities that can re-use the data and for what purposes the data can be used?	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant	Not relevant		Not relevant	Not relevant
	Publication of data	Do you plan to publish the data generated / collected in the pilot within its specific phase and if so, then where will you publish them?	Relevant	Relevant	Relevant	Relevant	Not relevant	Not relevant	Not relevant		Not relevant	Not relevant
DATA PRESERVATION AND ARCHIVING	*Archiving of data for preservation and long-term access*	How will the data produced within the pilot in its specific pilot phase be preserved for long-term access?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Not relevant
	Time period for data retention	How long the data should or could be retained?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Not relevant
	File formats of retained data	Please provide in what formats the data will be retained.	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant		Relevant	Not relevant
	Data archives	What type of data archives will be used to retain pilot generated/collected data?	To be determined	To be determined	To be determined	To be determined	To be determined	To be determined	To be determined		To be determined	Not relevant
	Long-term maintenance of data (systems and procedures)	Please provide the details on envisioned systems and procedures for long-term maintenance of data.	To be determined	To be determined	To be determined	To be determined	To be determined	To be determined	To be determined		To be determined	Not relevant

(continues including UCS-021 – UCS-030)

DMHP Questions	Use Case Scenario									
	UCS-021	UCS-022	UCS-023	UCS-024	UCS-025	UCS-026	UCS-027	UCS-028	UCS-029	UCS-030

DATA PRODUCTION AND STORAGE	*Types or categories of data generated/ collected*	What types of research data are collected or generated during a specific pilot phase? Who (or which entity) will be responsible for deciding what data is collected or generated?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Personal or non-personal data	Will the input or output data represent personal or non-personal data? What type of non-personal data will be collected at the pilot site in a specific pilot phase? What type of personal data will be collected?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Dummy/fake or real data	Will the data be dummy/fake/fictitious/synthetic or real?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Formats of the data	In which format will the data be collected (e.g., CSV, JSON, xls, PDF, ….)?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Reproducibility of data	Please provide the information for validation and reuse of data and indicate if the data are foreseen as open access	Relevant	Not relevant	Not relevant	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Data size	Please provide the information about the estimated size of data provided as input, as well as foreseen size of the data produced.	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Software tools for creating/processing /visualising data	Which application/ABB will be tested in your pilot in this specific phase? What aspects/functionalities of the applications will be tested at your specific pilot location in a specific phase? Besides the ABBs indicated as to be tested in your pilot phase, what other software tools will be used for creating/processing/visualising data?	GENERAL COMMENT: It is relevant, nevertheless, this question seems to be answered by the UCS themselves, as each of them specifically indicates the ABB and functionality which will be tested.									

	Use of pre-existing data	Will you use pre-existing data? Yes / No / Uncertain. If so, please indicate what pre-existing data will be used.	Relevant	Relevant	Relevant	Not relevant	To be determined	Relevant	Relevant	Relevant	Relevant	Relevant
	Data storage and backup strategies	Please indicate what storage and backup strategies will be adopted for input and output data.	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Purpose of data collection	Considering each type of data collected in the pilot phase, what is the purpose of their collection?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
ORGANISATION, DOCUMENTATION AND METADATA	*Standards for documentation of metadata*	What standards will be used for documentation and metadata (e.g., Digital Object Identifiers)? Is there a community standard for metadata sharing/integration?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Best practice/guidelines adopted for data management	Are there any best practices or guidelines which are foreseen to be applied in the context of organisation and documentation of the data and metadata in the pilot phases?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Tools for formatting data	What type of tools will you use to format data in the pilot phase?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Directory and file naming convention used.	What directory and file naming convention will be used? Will you provide clear version numbers?	Not relevant	Not relevant	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
DATA ACCESS	*Risks to data*	What main risks to data collected / produced during the pilot phase do you foresee? • Loss or destruction of data • Data breach • Loss of availability • Loss of integrity • Loss of confidentiality • Unauthorised alteration transmission and storage of data. Please provide any other major risks to data collected/produced at pilot sites	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant

	Risk management	Have you prepared a formal risk assessment addressing each of the major risks to data security and potential solutions? If so, please share further information. If no/uncertain, please explain why.	Relevant	To be determined	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Relevant
	Data access & requirements for access	Are there any concerns regarding access to your data? Yes / No	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Correct execution of the data access process	Please indicate a proper process which someone would need to take to access data collected/generated at pilot site during the pilot phase, as well as who is responsible for checking the correct execution of the access process. If data is confidential (e.g. personal data not already in the public domain, confidential business information or trade secrets), are there any appropriate security measures in place or any formal standards that you have to comply with?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Procedures to follow in the event of a data breach	Are there any specific data breach procedures which you foresee should be followed in the case of such an event?	Not sure	To be determined	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Relevant
DATA SHARING AND REUSE OF DATA	*Organization/labelling of Data for easy identification*	How will you organise or label the data to ensure that researchers may easily isolate fields of interest in their study?	To be determined	Not relevant	Relevant	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant
	Data Sharing & Audience for Data Sharing	Who can access data produced in the pilot and in a specific pilot phase?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Data Sharing Requirements	Are there any data sharing requirements which should be followed in the context of sharing the data produced/generated in the pilot in its specific pilot phase?	Relevant	Not relevant	Relevant	Not relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant

	Re-use of data	Will the data produced or generated in the pilot during its specific phase made re-usable or openly accessible? Will the data be reproducible (i.e. able to be copied)?	Relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Not relevant	Relevant	Relevant	Relevant
	Audience for re-use	Who will use the data during the pilot? Who will use it afterwards?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Restrictions on the re-use of data	Are there any restrictions regarding the entities that can re-use the data and for what purposes the data can be used?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Publication of data	Do you plan to publish the data generated / collected in the pilot within its specific phase and if so, then where will you publish them?	Not relevant	To be determined	Relevant	Not relevant	To be determined	Relevant	To be determined	Relevant	Relevant	Relevant
DATA PRESERVATION AND ARCHIVING	*Archiving of data for preservation and long-term access*	How will the data produced within the pilot in its specific pilot phase be preserved for long-term access?	Relevant	Relevant	Relevant			Relevant	Relevant	Relevant	Relevant	Relevant
	Time period for data retention	How long the data should or could be retained?	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	File formats of retained data	Please provide in what formats the data will be retained.	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant
	Data archives	What type of data archives will be used to retain pilot generated/collected data?	Relevant	To be determined	To be determined	Not Relevant	Relevant	To be determined	To be determined	To be determined	To be determined	To be determined
	Long-term maintenance of data (systems and procedures)	Please provide the details on envisioned systems and procedures for long-term maintenance of data.	Relevant	To be determined	To be determined	Relevant	Relevant	To be determined	To be determined	To be determined	To be determined	To be determined