D1.3 Annual Public Report II

Authors: Alexandra Garatzogianni, Alina Brockob (LUH), Gerrit Rosam, Michael Fribus (LUH)

Contractual Due Date: 31 December 2021 (M24)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871481

TRUSTS Trusted Secure Data Sharing Space

D1.3 Annual Public Report II

Document Summary Information

Grant Agreement No	871481	**Acronym**	TRUSTS
Full Title	TRUSTS Trusted Secure Data Sharing Space		
Start Date	01/01/2020	**Duration**	36 months
Project URL	https://trusts-data.eu/		
Deliverable	D1.3 Annual Public Report II		
Work Package	WP1		
Contractual due date	31/12/2021	**Actual submission date**	21/12/2021
Nature	Report	**Dissemination Level**	Public
Lead Beneficiary	LUH		
Responsible Authors	Alexandra Garatzogianni (LUH), Alina Brockob (LUH), Gerrit Rosam (LUH), Michael Fribus (LUH)		
Contributions from	Alexandra Garatzogianni (LUH), Gerrit Rosam (LUH), Alina Brockob (LUH), Michael Fribus (LUH), Anna Beer (LUH/TIB), Gianna Avgousti (EBOS), Ioannis Markopoulos (NOVA), Benjamin Heitmann (FhG), Kim Frieda Fidomski (FhG), Ilan Goldberg (EMC), Yuliya Miadzvetskaya (KUL), Lidia Dutkiewicz (KUL), Andreas Huber (G1), Bert Utermark (G1), Nina Popanton (DIO), Hannah Engel (DIO), Petr Knoth (RSA), Manos Paschalakis (REL), Evangelos Kotsifakos, Rosa Araujo (LST)		

© TRUSTS, 2021

Revision history (including peer reviewing & quality control)

Version	Issue Date	% Complete	Changes	Contributor(s)
v0.1	17/09/2021	20%	Initial Deliverable Structure	Alexandra Garatzogianni (LUH), Alina Brockob (LUH), Gerrit Rosam (LUH), Michael Fribus (LUH)
v0.2	26/11/2021	80%	Content	Alexandra Garatzogianni (LUH), Gerrit Rosam (LUH), Alina Brockob (LUH), Michael Fribus (LUH), Anna Beer (LUH/TIB), Gianna Avgousti (EBOS), Ioannis Markopoulos (NOVA), Benjamin Heitmann (FhG), Kim Frieda Fidomski (FhG), Ilan Goldberg (EMC), Yuliya Miadzvetskaya (KUL), Lidia Dutkiewicz (KUL), Andreas Huber (G1), Bert Utermark (G1), Nina Popanton (DIO), Hannah Engel (DIO), Petr Knoth (RSA), Manos Paschalakis (REL), Evangelos Kotsifakos, Rosa Araujo (LST)
v0.3	15/12/2021	90%	First Review	Nina Popanton (DIO)
v0.4	17/12/2021	95%	Second Review	Benjamin Heitmann (FhG)
v1.0	21/12/2021	100%	Final Version and Submission	Alexandra Garatzogianni (LUH), Gerrit Rosam (LUH), Michael Fribus (LUH)

© TRUSTS, 2021

Disclaimer

The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.
While the information contained in the documents is believed to be accurate, the author(s) or any other participant in the TRUSTS consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.
Neither the TRUSTS Consortium nor any of its members, their officers, employees or agents shall be responsible or liable for negligence or otherwise however in respect of any inaccuracy or omission herein.
Without derogating from the generality of the foregoing neither the TRUSTS Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

Copyright message

© TRUSTS, 2020-2022. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorised provided the source is acknowledged.

Table of Contents

© TRUSTS, 2021

© TRUSTS, 2021

List of Figures

© TRUSTS, 2021

List of Tables

© TRUSTS, 2021

Glossary of terms and abbreviations used

Abbreviation / Term	Description
BV	Business Validation
CRM	Customer Relationship Management
DL	Deep Learning
DMP	Data Management Plan
DoA	Description of Action
DS	Data Stewardship
DSM	Digital Single Market
E2E	End-to-End
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, and Reusable
FHE	Fully Homomorphic Encryption
FL	Federated Learning
FR	Functional Requirements
GA	Grant Agreement
GDPR	General Data Protection Regulation
HE	Homomorphic Encryption
HFL	Horizontal Federated Learning
IDS	International Data Space
IPR	Intellectual Property Rights
KPI	Key Performance Indicator
ML	Machine Learning
MPC	Multi-Party Computation
MVP	Minimum Viable Prototype

© TRUSTS, 2021

RAM	Reference Architecture Model
SAB	Stakeholder Advisory Board
SLA	Service-Level-Agreement
TL	Transfer Learning
TV	Technical validation
UC	Use Case
VFL	Vertical Federated Learning
WP	Work Package

© TRUSTS, 2021

1 Executive Summary

The objective of this **Annual Report II** is to report the project's progress in the second year of the TRUSTS Project. For Each Work Package (WP) and for each specific Lead there are reports on overall objectives, achieved progress in the first project year and the next steps. During the second project year (M13 – M24) the TRUSTS Consortium has achieved crucial goals in each WP. Some major highlights per WP include: In WP1, an all-encompassing, effective, forward-looking and collaborative project management of the whole TRUSTS project was ensured while enabling regular communication between the EC and the project consortium and monitoring potential risks. The final version of the TRUSTS architecture was defined in WP2, and the analysis of the worldwide data marketplace ecosystem was initiated. Key steps to realize the requirements and specifications of implementing the TRUSTS platform were realized in WP3. Moreover, concrete use cases for the use of smart contracts within the TRUSTS ecosystem were developed in this WP as well. In WP4, there was a focus on the development of risk analysis models and algorithms as well as on the implementation of corresponding modules for a ready-to-use application. The set-up of the test environment for the three use cases, as well as the planning and operation of the first execution phase of the use cases were undertaken in WP5. In WP6, an overview of the legal framework in order for TRUSTS to be compliant with the principles of research and ethics, and a set of legal and ethical requirements was set up with respect to potential legal and ethical obstacles. The results of the TRUSTS stakeholder landscape were transferred into suitable strategies for the TRUSTS stakeholder engagement plan in WP7. In WP8, the project's mission, vision, and achievements were actively communicated and disseminated to relevant TRUSTS stakeholder groups via various communication tools such as the website and social media channels. In WP9, Ethics deliverables have been resubmitted based on the requirements raised within the Ethics Check.

2 Introduction

This Deliverable (D1.3) is a report on the project's progress in the second year of the TRUSTS Project and consists of the following Sections. Section 3 contains reports of the objectives, the achieved progress and the next steps of each of the nine TRUSTS Work Packages. In Section 4 the objectives, the achieved progress and the next steps are specified for Scientific Lead, Technical Lead, Innovation Lead, Security Lead, Legal and Ethical Lead, Communication and Community Building Lead, Business and Exploitation Lead. The project's Data Management Plan (DMP) (D1.6) lists all relevant information on current and planned data management activities. It will be regularly updated to reflect the development and progress of the project. Section 5 contains a first update of the DMP. Conclusions are outlined in the last Section 6.

Mapping Project's Outputs:

Purpose of this section is to map TRUSTS Grant Agreement (GA) commitments, both within the formal Deliverable and Task description, against the project's respective outputs and work performed.

© TRUSTS, 2021

Table 1: Adherence to TRUSTS GA Deliverable & Tasks Descriptions

Task		Respective Document Chapter(s)	Justification
T1.1 Project Management	This task deals with all necessary project management tools, mechanisms and structures for the high quality, efficient and timely administrative coordination of the project. It incorporates Administration Management activities, including procedures and guidelines for activity planning and monitoring, cost and time management, submission of periodic progress reports and cost statements, preparation of annual review reports, review presentations, and timely submission of deliverables to the Commission. LUH will be responsible for the day-to-day coordination of project related activities and tasks, as well as the administrative management of the project; contributions will be made by all the partners.	Section 3 – Section 5	Section 3: Progress within Work Packages (WP) Section 4: Progress within specific Leads Section 5: Data Management Plan
Deliverable			
D1.3 Annual Public Report II Report on the project's progress, targeting the general public. The report will focus on the impact of the conducted work.			

3 Progress within the Work Packages

3.1 WP1 Project Management

3.1.1 Objectives

The Project Management WP serves the objective to ensure the timely, successful and impactful delivery of the project's results in compliance with the EC regulations and the H2020 framework. This is achieved via the hands-on and continuous monitoring of the implementation and completion of the project's tasks, activities, milestones and deliverables, safeguarding thus their proper and timely development according to the Description of Action (DoA) and the project's work-plan, while ensuring the smooth and efficient collaboration among the consortium partners.

The activities of the project management WP focuses on establishing tasks, providing thus guidance and direction to achieve the goals of the H2020 TRUSTS project, ensuring continuous, proactive communication with the EC, establishing efficient means of communication and document exchange between partners, ensuring transparency at all levels and in terms of reporting by establishing appropriate report structures and procedures, conducting quality assurance activities and performing risk analysis tasks, coordinating the organisation of project meetings and other possible participatory events where the project could be presented and ensuring that project objectives are realised within set time, quality and budget.

3.1.2 Progress achieved

T1.1 Project Management

In the scope of this task, LUH ensured a **high-quality and efficient administrative coordination** of the TRUSTS project. This task included responsibilities such as a hands-on consistent **Consortium Project Management**, the organization and lead of **regular PM telcos** (incl. agenda and minutes preparation), continuous and proactive **costs and time management**, the identification of **possible technological and managerial issues** including **mitigation plans** to tackle these, and the preparation of annual review reports and review presentations. The **TRUSTS Review** was successfully organized in M21, on 14th of September, 2021. In collaboration with DIO, G1 and IDSA, LUH has successfully prepared the TRUSTS **Stakeholder Advisory Board (SAB)** which was kicked-off in M18.

Moreover, LUH was responsible for the review, the **quality check** and the timely **submission** of high-quality **deliverables** to the EC. During 2021, 11 deliverables have been submitted and one deliverable has been re-submitted to the EC portal. This means that as of now (M24) **48 out of 70 deliverables have been successfully finalized and submitted** to the EC.

An overview of the deliverables that have been submitted during the period M13–M24 can be seen in the table below.

© TRUSTS, 2021

Table 2: Deliverable Status Overview (M13-M24)

Del.	Title	Lead	Due	Status
D3.9	Platform Status Report I	FhG	M12	**SUBMITTED**
D5.1	Pilot planning and operational management reports I	EBOS	M14	**SUBMITTED**
D6.1	Research Ethics	KUL	M14	**SUBMITTED**
D2.1	Definition and analysis of the EU and worldwide data market trends and industrial needs for growth	IDSA	M18	**SUBMITTED**
D3.7	Data Governance, TRUSTS Knowledge Graph I	SWC	M18	**SUBMITTED**
D3.12	Profiles and Brokerage I	KNOW	M18	**SUBMITTED**
D4.1	Algorithms for Privacy-Preserving Data Analytics	KNOW	M18	**SUBMITTED**
D7.1	Sustainable business model for TRUSTS data marketplace I	TUD	M18	**SUBMITTED**
D7.3	Communities engagement strategy	IDSA	M18	**SUBMITTED**
D7.4	Supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship I	G1	M18	**SUBMITTED**
D7.7	Business plan and Implementation action plan I	LST	M18	**SUBMITTED**
D7.9	Innovation Impact Assurance I	G1	M18	**SUBMITTED**
D8.6	Concept for training and capacity building programme	REL	M18	**SUBMITTED**
D1.3	Annual Public Report II	LUH	M24	**SUBMITTED**
D2.3	Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II	NOVA	M24	**SUBMITTED**
D2.5	Methodologies for the technological/business validation of use case results II	EBOS	M24	**SUBMITTED**
D2.7	Architecture design and technical specifications document II	FhG	M24	**SUBMITTED**
D3.5	Data Marketplaces with Interoperability Solution II	RSA	M24	**SUBMITTED**
D3.10	Platform Status Report II	FhG	M24	**SUBMITTED**
D5.4	Actual field trials of use case 1. v.1	EBOS	M24	**SUBMITTED**
D5.6	Actual field trials of use case 2 v.1	NOVA	M24	**SUBMITTED**

© TRUSTS, 2021

D5.8	Actual field trials of use case 3 v.1	REL	M24	**SUBMITTED**
D5.10	Performance evaluation and lessons learned report I	NOVA	M24	**SUBMITTED**
D8.4	Annual Dissemination Report II	DIO	M24	**SUBMITTED**

T1.2 Technical & Quality Assurance and Risk Management

For this task, appropriate **mechanisms and processes** have been established **to maintain overall quality** in all WPs of TRUSTS. In particular, WP1 oversees an extensive review process of all TRUSTS deliverables. Thus, for the reporting of deliverables and for managing the overall quality of the deliverables, LUH has set up a dedicated **review process for the deliverables** which can be seen below.

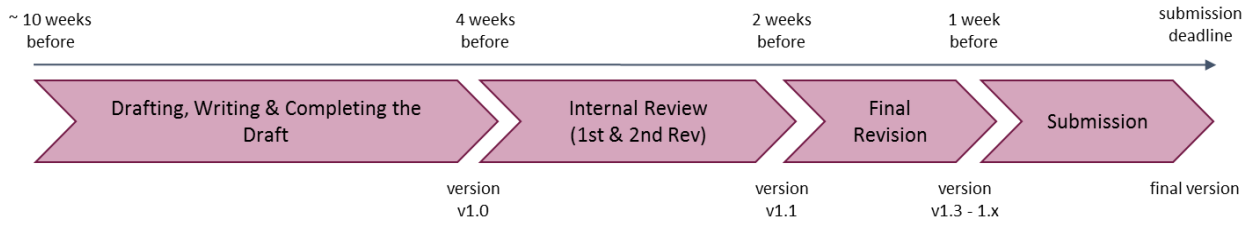

Figure 1: Deliverable Drafting, Review and Submission Process

The content of the deliverables of each TRUSTS WP is checked against 'Deliverable Quality Indicators' such as format, readability and consistency (e.g. contains right information, avoids redundant information, consistent with previous deliverables, etc.). Also, two Consortium partners were assigned to act as peer reviewers for each deliverable.

LUH has **identified and monitored potential PM Risks** and has successfully developed respective **mitigation plans** and established a **Risk Impact Assessment** for Risks from TRUSTS' DoA, which is discussed regularly in terms of Project Management Board (PMB) Telcos and updated when needed.

T1.3 Project Reporting & Communication

LUH ensured the implementation of regular reportings, milestone reviews and the Mid-Term Review (M18) as stated in the DoW. Thus, The **Project Management Report** as well as the **Financial Report** were finalized and timely submitted in M18. Moreover, a setup of various, **regular PM calls** were organized by LUH in the period M13-M18 such as the WP1 PM Executive Board Telcos, PMB Telcos, the two Plenaries in M13 and M18, among others. As the Coordinator of TRUSTS, LUH functioned as a **contact point** between the EC and the TRUSTS Consortium. Thus, LUH ensured the **regular communication between the EC and the project consortium** as well as the **communication within the TRUSTS consortium**.

An overview of the **Project Reporting & Communication** activities can be seen in this table:

© TRUSTS, 2021

Table 3: Project Communication Formats

Project Reporting and Communication activities		
Types of Telcos	**Frequency**	**Purpose**
Executive Board PM WP1 Telcos	monthly	A call that provides oversight and an update of WP1, as well as the other WPs in TRUSTS. It gives participants the opportunity to learn about the progress, status quo and challenges of each WP.
Plenary	bi-annual	A deep dive into the project in which partners discuss the progress and proactively work on solutions for current challenges.
EC Review	after 18 months	A periodic conference with the purpose to report the progress and achievements of TRUSTS and its WPs to the EC. In total, there are two EC Reviews that take place during the TRUSTS project.
Project Management Board Telcos (PMBs)	bi-/ tri-monthly	A high-level conference to discuss the strategic development of TRUSTS with focus on specific tech / business-related / organizational tasks and challenges.

3.1.3 Next Steps

Looking back at the second project year 2021 (M13-M24), it can be summarized that WP1 and the TRUSTS project as a whole proceeded as planned. In the next project year 2022, the ongoing PM work will be continued and finalized in M36. It will be ensured that the **high quality work** will be maintained in all WPs and the project as a whole. All necessary steps will be made so that all project outcomes are reached in time and that the corresponding budget will be used accordingly. The TRUSTS partners will be guided and supported by LUH while ensuring that potential items will be addressed proactively and forward-looking. The conditions that are indispensable to meet the project's targets will be safeguarded by the project coordinator, as specified in the GA. LUH will manage the **quality control and submission of the remaining 22** (out of 70) **deliverables**. In particular, the next **D1.4 Annual Public Report III** will be finalized and submitted on time by LUH. The final results of TRUSTS will be presented in the next **PM Report** as well as **TRUSTS Review**, which will both take place in M36.

© TRUSTS, 2021

3.2 WP2 Requirements Elicitation & Specification

3.2.1 Objectives

The overall objectives of WP2 as defined in the DoA are:

- to analyse the EU and worldwide challenges and trends and to define the requirements for the provision of a multi, concurrent and cross-domain, secure and scalable end-to-end (E2E) data marketplace service.
- to define detailed and functional industry specifications appropriate for a data marketplace linked to specific target KPIs considering and bridging the vertical user point of view (PoV) with the analytics/solution provider PoV and the data marketplace platform provider PoV.
- to produce a set of KPIs and methodologies to enable:

  (a) the technological and Business Validation (BV) of the E2E data marketplace service and associated control and management within and across verticals;

  (b) the definition of the test reports format, parameters, test points, and benchmarking of the results for a unified and reliable outcome.

During the second year of the project, WP2 focused on:

- Initiating the analysis of the worldwide data marketplace ecosystem in terms of status, markets, trends, success and failure stories.
- To set up the process for the second iteration and finalisation of the Functional Requirements (FR). The analysis of a wide variety of sources (updated online surveys, stakeholders interviews, TRUSTS deliverables in market analysis – D2.1, architecture – D2.6, Business and remuneration models – D7.1, D7.4, and use case trials) in order to define the final TRUSTS FR.
- The update and evolution of the trials evaluation testing methodology and the respective business evaluation methodology. The methodology required acceptance test procedures for conducting both the technological and BVs of the UC's considering the associated service management. The objective is to validate the three UC's – business wise – and develop business plans for the UCs with the highest commercial potential.
- The definition of the final version of the TRUSTS platform architecture. The architecture represents the conceptual foundation for the implementation of the TRUSTS platform, and therefore reaching a consensus on the architecture enables all project partners with a technical view to agree on the most important abstract decisions, before realising them in their implementation. In addition, the architecture also allows the project partners with a non-technical view to contribute with cross-cutting requirements of strategic importance, such as having future proof characteristics e.g. compliance to GAIA-X concepts.

In the second year WP2 produced the following deliverables:

- **D2.1**: Definition and analysis of the EU and worldwide data market trends and industrial needs for growth [M18]
- **D2.3**: Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II [M24]
- **D2.5**: Methodologies for the technological validation/BV of use case results II [M24]
- **D2.6, D2.7**: Architecture design and technical specifications document I, II [M17, M24]

The work in WP2 was organised in 4 tasks as illustrated in the figure below, and concluded in M24, December 2021:

© TRUSTS, 2021

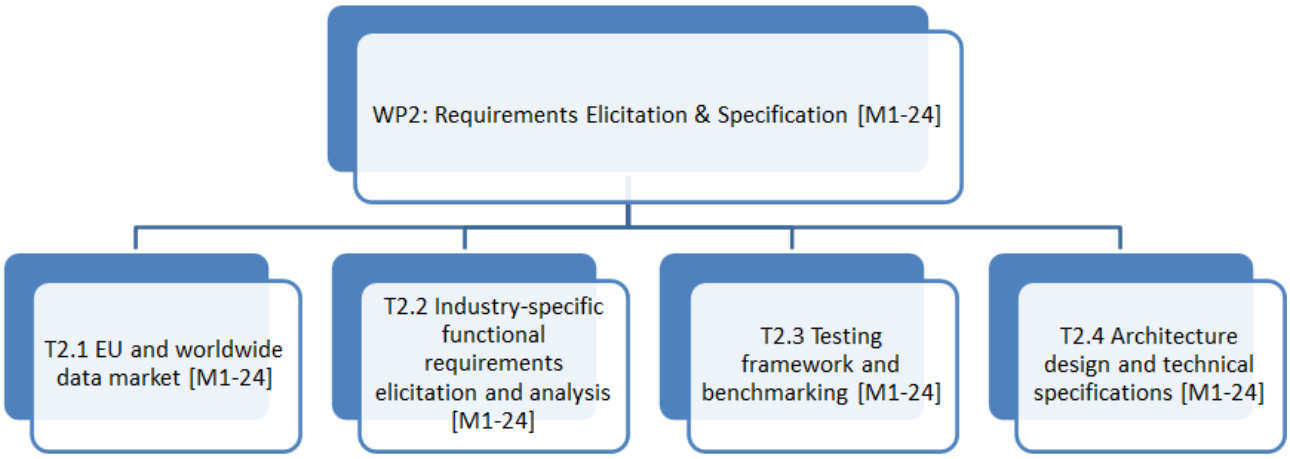

Figure 2: Tasks in Work Package 2

3.2.2 Progress achieved

This WP has progressed in all fronts demonstrating tangible results from all four tasks which are detailed within this Chapter.

Task 2.1 EU and worldwide data market

In this task, led by IDSA, a study that depicts several facets of the environment data marketplaces are embedded in and that are of relevance for the TRUSTS project was conducted. The aim of this task was to provide an overview on current states and relevant trends to ensure that the project's results are targeting real market needs and working with up-to-date requirements. Within the study, an analysis of the academic landscape on data marketplaces took place, delivering among other things a framework to classify and position TRUSTS results as depicted in the following graphic:

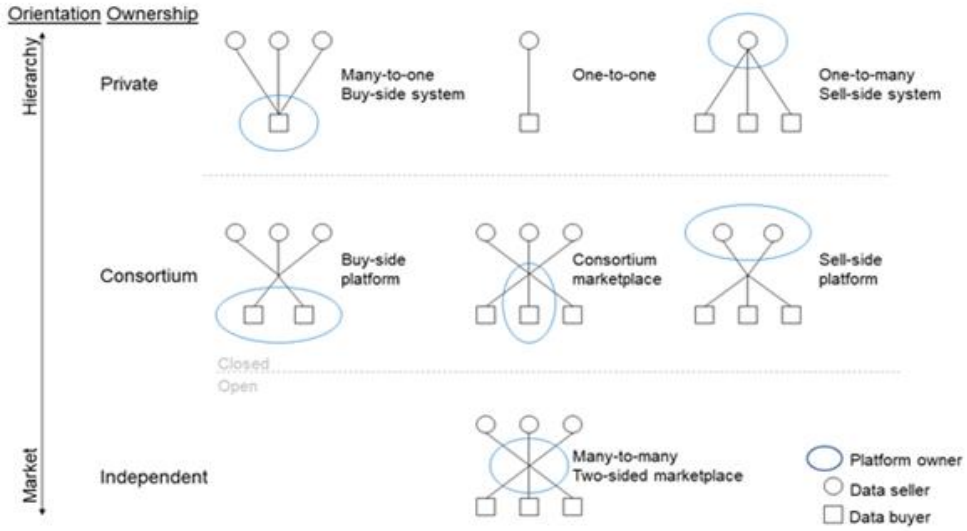

Figure 3: Data Marketplaces categorized by Orientation and Ownership

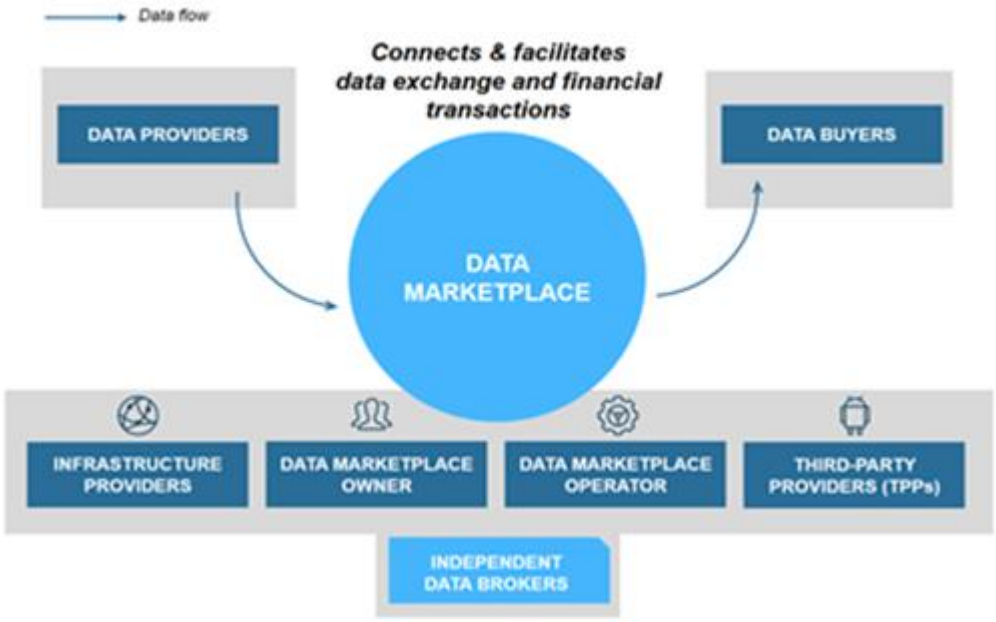

Figure 4: Facilitation of Data Exchange and Financial Transactions

A definition of a data marketplace as illustrated in the figure above was elaborated. Additionally, an analysis of the macroenvironment of data marketplaces, examining and summarizing the current circumstances and trends in the following areas that affect and are affected by data marketplaces: political, economic, social, technological, legal, and ecological areas were conducted. Here, relevant developments such as the Gaia-X initiative, examples of implemented business models archetypes (relevant to WP7), current formats of data marketplace system architectures, features and a list of relevant and used standards were identified. Then we moved to the microenvironment of data marketplaces, by analysing the direct competitive environment, using Porters' Five Forces Analysis and discussed the relevance of data market federators. For each of the sections we examined we derived suitable recommendations valid not only for TRUSTS but also upcoming endeavors in this regard. These recommendations will be provided to the respective tasks, in order to inform the analysis of requirements at T2.2 and establishment of competitive and/or complementary specifications at T2.4.

For some of the areas, external experts on data marketplaces (predominantly from industry) were invited with an attempt to gather additional insights on current trends and issues, in the form of a "world café" workshop. The topics were: environmental, social and technical aspects of data marketplaces, but also business models for data marketplaces and data sharing versus data trading. Here, a collaboration with WP7, T7.2 (Developing and structuring the platform engagement) and WP8 (Dissemination, Communication & Community Building) took place in order to organize this event that took place digitally in March 2021.

The final deliverable D2.1 was submitted in M18. In the following months, Task 2.1 ensured that all final recommendations were updated and provided to the respective WPs. Also, circulated, communicated and promoted the results of the study together with T7.2 and WP8 to external stakeholders in order to foster the community around TRUSTS.

© TRUSTS, 2021

Task 2.2 Industry-specific functional requirements elicitation and analysis

The aim of this task was to capture requirements from a wide range of domain stakeholders in order to produce recommendations for the E2E data marketplace operation for, but not limited to, the telecom and financial sectors. To achieve this, a systematic methodology was adopted, analysing requirements from a wide variety of information channels, sources and stakeholders, i.e.:

1. The use of an updated electronic **survey**.

   The majority of participants in the survey identified themselves as business or technical drivers at all levels of management hierarchy, with many years of experience in the field and an understanding of the buying/selling processes in the organization. Regarding their level of management, a considerable proportion of participants clarified their role as administrative officers (31%), closely followed by researchers (28%). Operating officers and university professors each correspond to 17%, while executive officers each represent 8% of the participants.

   Indicative requirements resulted from the survey responses analysis:

   "Responders envisage TRUSTS as a One-stop-shop online service for buying and selling data"

2. Key stakeholders' **interviews**.

   Significant effort was made to interview executives (director level and above) external to the project. This was achieved, to a large extent, and mainly the respective interview impacted TRUSTS requirements.

   Indicative finding from the interviews:

   "There is a lot of potential but in order to be successful one has to access sources of really big data respecting security and IPRs. Business may start from vertical markets or big industries as clients and their ecosystem. A business alignment with such industrial partners could be beneficial."

3. The in-depth analysis of the EU and worldwide data market trends and industrial needs for growth (delivered in D2.1). The figure below illustrates the key marketplace features identified in D2.1:

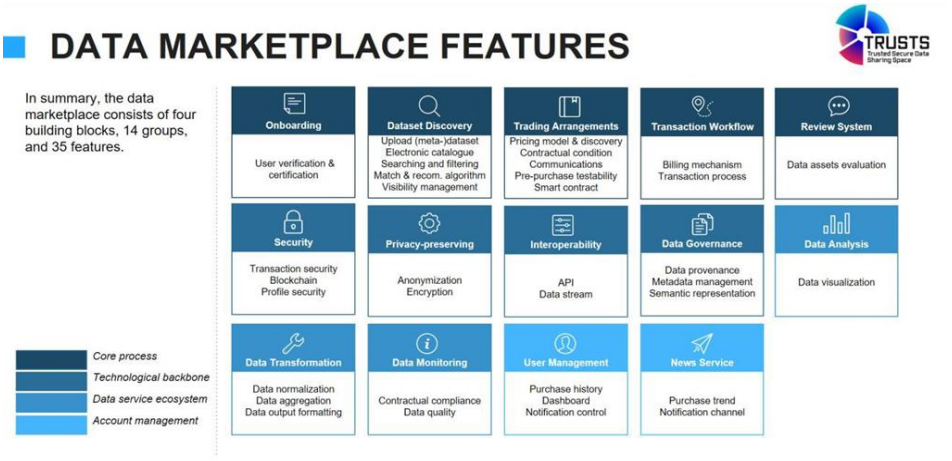

Figure 5: Data Marketplace Features

4. The analysis on the state-of-the-art business processes and models (delivered in D7.1).

   Indicative requirement: “For TRUSTS to become an ecosystem facilitator, it is required to create a business and commercial plan on defining a series of actions that enable data governance models and other framework conditions allowing companies and individuals to avoid the negative externalities of proprietary industrial platforms (supply-driven approach, lower level of control on proprietary data, centralized data governance and technical architecture). Attracting an ever-increasing number of companies and achieving critical mass would be fundamental for TRUSTS to become recognized and successful and a wide range of domain actors.”

5. The analysis of the supporting mechanisms for Intellectual Property Rights (IPR) Protection and Data Stewardship (DS) (elaborated in D7.4)

   Indicative requirement: “TRUSTS may use mechanisms defined in International Data Space Reference Architecture to support the IPR protection”

6. Architectural considerations for the Use Cases implementation and Business Model realization (provided in D2.6)

   This deliverable resulted in the refinement of 5 FR defined in the Deliverable D2.2.

The methodology adopted by T2.2 towards producing the updated FR is illustrated in the following figure:

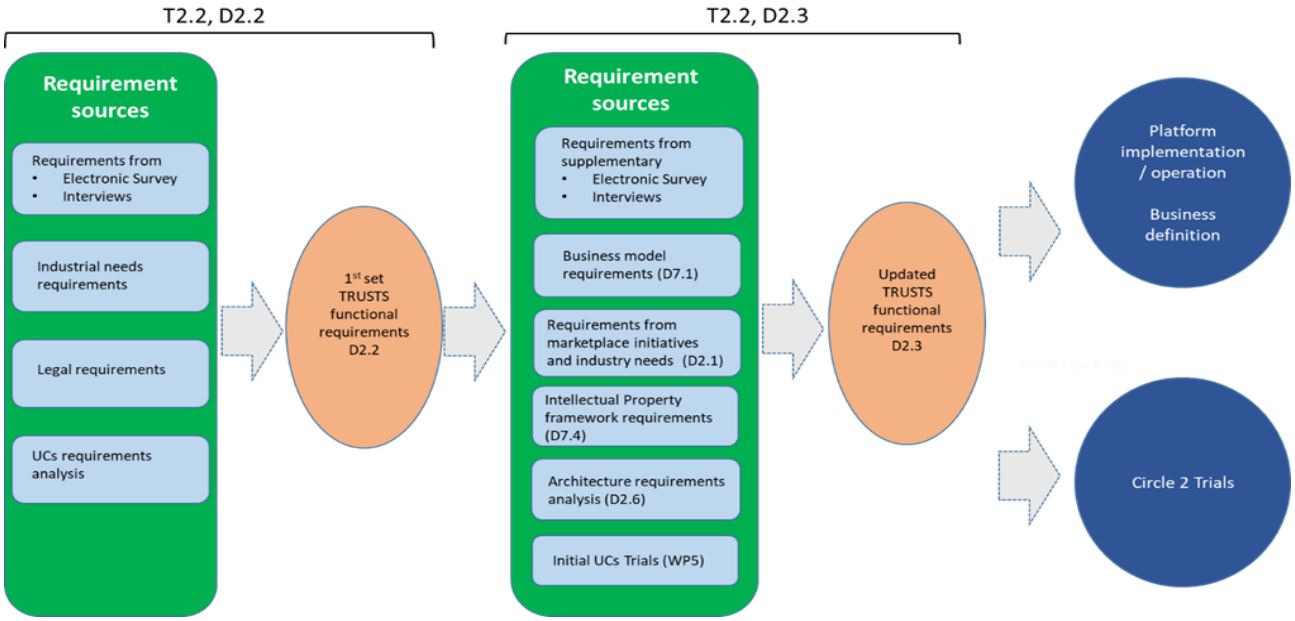

Figure 6: Methodology for producing updated Functional Requirements

All requirements sources are analysed for individual requirements and their justification.

- The above mentioned requirements drive the definition of the updated set of FR, which will be used for the implementation of the TRUSTS platform, as well as the evolving operational processes design.

© TRUSTS, 2021

- The combined updated set requirements were evaluated taking also into account the outcome of Circle 1 Trials.

The analysis resulted in a list of 51 updated FR augmenting and when necessary modifying the D2.2 FR list. FR are categorised as follows:

- Datasets and services onboarding functionality and processes
- Intelligent data/service exploration and correlation functionality and processes
- Purchasing transactions and billing
- (Meta-) Data Governance
- Data as a Service (DaaS) and Subscribers management
- Data protection
- Advanced data analysis based on Machine Learning (ML)
- Trusted and legitimate data flows

The updated FR were presented in an internal TRUSTS consortium workshop in November 2021 and subsequently all partners were asked to provide their comments and evaluation. Following the respective round of comments (resulted on the 15th of November) all FR were accepted as guiding points to the TRUSTS platform.

The analysis of the updated FR is reported in the Deliverable D2.3.

Task 2.3 Testing framework and benchmarking

Task 2.3 was led by EBOS and its aim was to define the methodology and toolset for the analysis and validation of the data marketplace technologies and the three Use Cases (UCs) implemented in TRUSTS.

Task 2.3 worked closely with the partners of Task 2.2 and WP5 “Demonstration of the TRUSTS Platform in three business-oriented Use Cases” to define the detailed scenarios to be trailed in the TRUSTS environment that were continuously updated and improved supporting the UC trials. Task 2.3 also supported the test cases, while in order to measure the functionality and performance, in collaboration with Task 2.2 defined the 44 FR reported in D2.2.

The methodologies for the technological and business validation of the TRUSTS Platform within and across each UC were also defined and documented in D2.4 that was submitted in June 2020 (M6). An updated version including the revised templates and the consolidated results that fed to WP5, WP7 and WP3 accordingly, are documented and in parallel submitted by December 2021 to the D2.5 report.

- The Lean Start-Up Methodology was defined with respect to the BV.
- The procedure of the test-driven development methodology (taking specifics from Task 2.1) and toolset for the analysis of the data marketplace technologies and the UCs for the purposes of the technological validation that are held during the life period of the project, were defined.
- These in order to receive end-users’ feedback and to set the metrics and parameters and also to present the KPI validation so as to enable the project to focus on the lessons learned within.
- Task 2.3 also focused on the methodology to be followed during the UC trials and what needs to be tested, which allowed the validation and evaluation of the functionality and performance of the marketplace to later deliver outputs that have commercial value and potential.

© TRUSTS, 2021

A Test Case Validation Toolset (See Figure 7), and a number of business and technical validation templates (Figures 8 and 9) were defined and later used for the evaluation of TRUSTS and the offered services – as a whole – from each UC (conducting both the technological and business validation of the UCs).

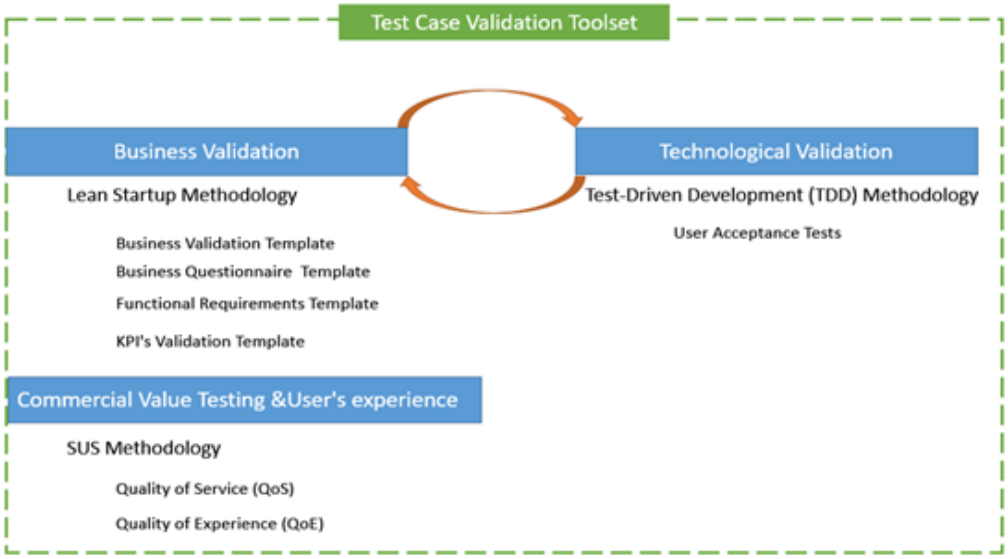

Figure 7: Test Case Validation Toolset

Background *Please provide a textual description of the business process and context surrounding the UC.*	
What is the general context of the UC? (describe the Organisation / business situation) *Under what circumstances does the UC arise?* *How often?* *Other information?*	
Describe the Personas *Please describe ALL personas who are **directly** impacted by the UC* *Describe each persona of the TRUSTS (Consumer? Org/Business operations? Technology? Etc.);* *Please be as specific and detailed as possible about exactly what each persona does.*	
*Describe the **end user personas** (e.g. different types of consumers; operators in a data marketplace?)*	
Persona Name	*Persona Role*
*Describe the **application provider(s)** (who builds and supports the application?)*	
Persona Name	*Persona Role*
*Describe **other actors** directly involve/impacted by the UC?*	
Persona Name	*Persona Role*
End user personas	
Describe the Problem *Describe in detail the problems that each persona/stakeholder currently experience (AS-IS today before TRUSTS)* **Personas** (who exactly?) experience this **problem** (what exactly?) when doing this **task** (when does it occur?) OR **Personas** (who exactly?) experience this **problem** (what exactly?) because of this **constraint** or limitation (when does it occur?)	

Figure 8: Business Validation template (Part 1)

UAT Scope			
UAT - In Scope		**UAT - Out of Scope**	
In Scope *List features that are tested.*		Out of Scope *List features that are not tested.*	
UAT Assumptions and Constraints			
UAT Assumptions			
Assumption *List the UAT assumptions/expectations.*			
UAT Constraints			
Constraint *List the UAT constraints/limitations.*			
UAT Risks			
Description	**Probability** **High | Medium | Low**	**Impact** **High | Medium | Low**	**Mitigation**
Risk *List the risks of UAT.*	*How likely is the risk to occur?*	*What is the impact of the risk on the UAT?*	*Steps to avoid the risk.*
Add more rows if needed.			
UAT Team Roles & Responsibilities			
Name	**Roles**	**Responsibilities**	
List names of people involved in testing.	*i.e., UC leader, stakeholder, observer, technical prs etc.*		
List names of people involved in testing.			
Add more rows if needed.			
UAT Entry Criteria			
Criteria			
Entry Criteria *Factors that must be present to enable the start of the UAT.* ***Example**: Testing environment/VMs/nodes/data available etc.*			
UAT Requirements-Based Test Cases			
Test Cases			
Test Case 1 *Identify the test cases along with the expected results.* *Test Procedure:* *Login with a corporate user account.* *Expected Results:* *An error will be displayed for the wrong credentials.*			
Test Case 2 *Identify the test cases along with the expected results.* *Test Procedure:* *Expected Results:*			

Figure 9: Business Validation template (Part 2)

UAT Test Results			
Test Cases	**Pass/Fail**	**Tested By**	**Date Tested**
Test Case 1 *Name the test case.:* *Test Procedure:* *Expected Results:*			
Test Case 2 *Name the test case.:* *Test Procedure:* *Expected Results:*			
Test Case 3 *Name the test case.:* *Test Procedure:* *Expected Results:*			
Add more rows if needed.			
Addendums & Appendices			
Include any additional documents or link to screenshots/video to support the above			

Figure 10: Technical Validation template

The technical and business validation go hand in hand allowing us to capture the project objectives and satisfy end-user needs via the overall data marketplace evaluation over the 3 UCs.

As shown in the Figure 10 below:

3 periods of BV were defined:

- First Business Validation M7 – M10 (July 2020 – October 2020)
- Second Business Validation M23 – M24 (November 2021 – December 2021)
- Third Business Validation M30 – M33 (June 2022 – September 2022)

2 periods of Technological Validation were defined:

- 1st Technological Validation M19 – M22 (July 2021 – October 2021)
- 2nd Technological Validation M25 – M29 (January 2022 – May 2022)

The major requirement here is for the project to deliver outputs that have commercial value and potential. The corresponding outputs allow us to identify the use cases that have the highest commercialization potential in order to progress.

The **first BV** was performed between July and Oct 2020 as per the Gantt chart (see Figure 10 below), created under Task 2.3 with internal milestones matching with the project's milestones as well. The first BV outputs to the related WP3, WP4, WP5 and WP7 were:

- A detailed collection of business information about the UCs including the description of the problem (before TRUSTS) and the expected benefit (after TRUSTS), different personas, their role and who are directly impacted by the UC.
- A definition of a number of scenarios to be executed on each UC along with the expected results and a mapping of requirements and functionalities for each scenario.
- Finally, an attempt to define what are the expected (required or nice to have) functionalities provided by the TRUSTS data marketplace which will benefit at a business level the involved parties of each UC.
- And the revised KPIs per UC by giving some more information about them (including the Baseline value, target value by M36, calculation and validation method).

Validation Type	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18
	Jan-20	Feb-20	Mar-20	Apr-20	May-20	Jun-20	Jul-20	Aug-20	Sep-20	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21
						MS1						MS2						MS3
WP2 - T2.3																		
Business Validation (BV)							1st BV											
UC trials											Preparatory activities - planning, setup of UC's + deployment					1st phase		
Deliverables						D2.4 (eBOS)												
Validation Type	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36
	Jul-21	Aug-21	Sept-21	Oct-21	Nov-21	Dec-21	Jan-22	Feb-22	Mar-22	Apr-22	May-22	Jun-22	Jul-22	Aug-22	Sep-22	Oct-22	Nov-22	Dec-22
							MS4											MS5
WP2 - T2.3																		
Business Validation (BV)					2nd BV							3rd BV						
Technological Validation (TV)	1st TV						2nd TV											
UC trials				Preparatory activities - planning, setup of UC's + deployment			2nd phase											
Deliverables						D2.5 (eBOS)												

Figure 11: Task 2.3 Gantt chart

The **first technical validation** was performed by the UC participants during the first trial period (May – November 2021), aligned with Milestone's timeline since it is initiated right after Milestone 3 "First Pilot Deployment" (M18). It allowed the test and validation outcome of the existing technical implementation through predefined scenarios and templates, allowing the validation of the architectural framework and technical specifications (T2.4) along with the work under the T3.5 "Initial Platform and integration", assessing its correct functioning according to its functional and technical requirements. The objective was to validate the three UC's technical wise with technical and interoperability testing issues since before transferring the technology to the market, it first must be validated. This validation gave feedback to T2.4 and WP3, and it is further elaborated in D2.5 set to be submitted December 2021 with a completed set of the technical templates used available.

Following the **first business** and the **first technical validation** that were performed by the UC participants during the first trial period, allowing them to check and validate the outcome of the existing technical implementation through predefined scenarios and document the results, in the report produced under this Task, D2.5 submitted in parallel by December 2021.

This is aligned with the first phase of the trials and the WP5 collaboration finished by December 2021 (M24). Task 2.3 revised the methodologies already defined in D2.4 and updated the validation toolsets while acknowledging the – so far – feedback from the UCs and their first phase of the trials. Therefore T2.3 assisted and will keep assisting the three UCs using these templates and the evaluation methodologies developed during the lifetime of the project and more specifically the TRUSTS trials.

The aim is to systematically assess the input from all involved parties in order to fulfil the objective of T2.3, by validating the three UC's business and technical wise and develop business plans with the highest commercial potential.

Moving on, the **second BV**, is set to be performed within November and December 2021 (M23-24), starting prior the second cycle of the trials, and finalised by the end of this task and WP. It will be performed by the UC leaders as well based on the first cycle outcome as well as the plan for the second cycle. The outputs of the second BV will also give input to the D5.2 "*Pilot planning and operational management reports II*" (due January 2022) as per the planning of the second cycle of the project's trials starting January 2022, as well as to the respective deliverables of each UC reporting on the actual field trials and environment, concluding in August 2022. The UCs input is set to be collected by December 2021 and the actual outcome will be reported in the closing deliverable of WP5 concluding the final cycle of the UCs trials.

The **second Technical Validation** is planned to be performed between January 2022 (M25) to May 2022 (M29), allowing the validation of the Marketplace and the provided services during the second set of the UC trials by utilizing the defined test procedures and the reporting structure, and validation of results regarding technology. This validation will be aligned with the milestone's timeline since it is initiated right after Milestone 4 "End of second period" (M24) and performed by the UC participants during the second set of the UCs trial period, allowing them to check and validate the outcome of the technical implementation through predefined scenarios and document the results using the above templates. This last round of technical validation will also evaluate the complete environment from a technical, performance, expandability (e.g., federation etc.) point of view and define the quality of the implementation. The output will be an input back to WP3 and WP4 for the refinement of the implemented solution (marketplace).

The **third BV** is set to be performed from June 2022 – September 2022 (M30 to M33), allowing the evaluation of the complete environment from a performance and business point of view, via the measurement of the UCs KPIs and validation of their results to define the gap towards commercializing the

© TRUSTS, 2021

environment. This last round of BV will be performed again by the UC participants where the output of this final BV shall be an input to WP7.

Under this Task 2.3, the first version of the two reports produced defining the methodologies for the technological and BV of the TRUSTS Platform within and across each vertical UC was submitted in June 2020 (M6) titled: D2.4: Methodologies for the technological/business validation of use case results I. The second and final version of the report regarding the revised methodologies and results for the technological and BV of the TRUSTS UCs during the first cycle of the TRUSTS trials, is set to be submitted by the end of this year December 2021 (M24) as "D2.5 Methodologies for the technological/business validation of use case results II".

As the corresponding WP and Task 2.3 conclude in December 2021, the remaining validation efforts will be performed by the UC participants as planned and illustrated in Figure 10. While finishing the second demonstration phase by August 2022 (M32) the third and final BV will be completed by September 2022 (M33). These validations will give feedback to WP3, WP5 and WP7 accordingly in regards to the TRUSTS Platform.

Task 2.4 Architecture design and technical specifications

The work in T2.4 during the second year focused on the second iteration of the TRUSTS architecture and the production of D2.7: "Architecture design and technical specifications document II". For the second version of the architecture, the focus was on iterating the architecture based on feedback from the use case partners and from the non-technical project partners. The project partners with a technical perspective have refined the architecture and documented additional technical specifications based on their implementation experiences.

The architecture design of the TRUSTS Platform represents the blueprint for the technical results of the TRUSTS project. As such, it also represents the foundation for instances of the TRUSTS Platform which will be provided by one or more TRUSTS operators after the duration of the project. The technical specifications provide the details which are required by technical experts in order to instantiate the platform infrastructure and build on top of it or to extend it with their own components, services and applications.

The architecture is innovative beyond the state-of-the-art, as not just data sets can be traded, but also access to services and applications while maintaining security and privacy of all involved participants. The architecture of the TRUSTS Platform has to accommodate the requirements and priorities of many different stakeholders inside and outside of the project. In order to accommodate this, the participants of the task went through the following **process**:

Technical requirements for the architecture from external parties: First, we collected architectural requirements from relevant external parties. We collected requirements for the architecture from the initiatives on which the TRUSTS Platform is based, i.e., Data Market Austria (DMA) and International Data Spaces (IDS). In addition, we collected architectural requirements from the Gaia-X initiative, as future compatibility with Gaia-X is of strategic importance to the TRUSTS Platform. We expect Gaia-X to set important impulses for the data economy in Europa by, e.g., communicating with important groups of stakeholders to set the agenda and by setting standards for technical and organizational issues, such as certification.

Technical requirements for the architecture from parties within the project: Then we collected architectural requirements from the different technical participants within the project, grouped by areas of concern. We

© TRUSTS, 2021

collected such requirements related to: smart contracts; interoperability of data marketplaces; data governance; platform development and integration; brokerage and profiles for users and corporates; privacy enhancing technologies; anonymization and de-anonymization; as well as from the usage of the CKAN data portal software.

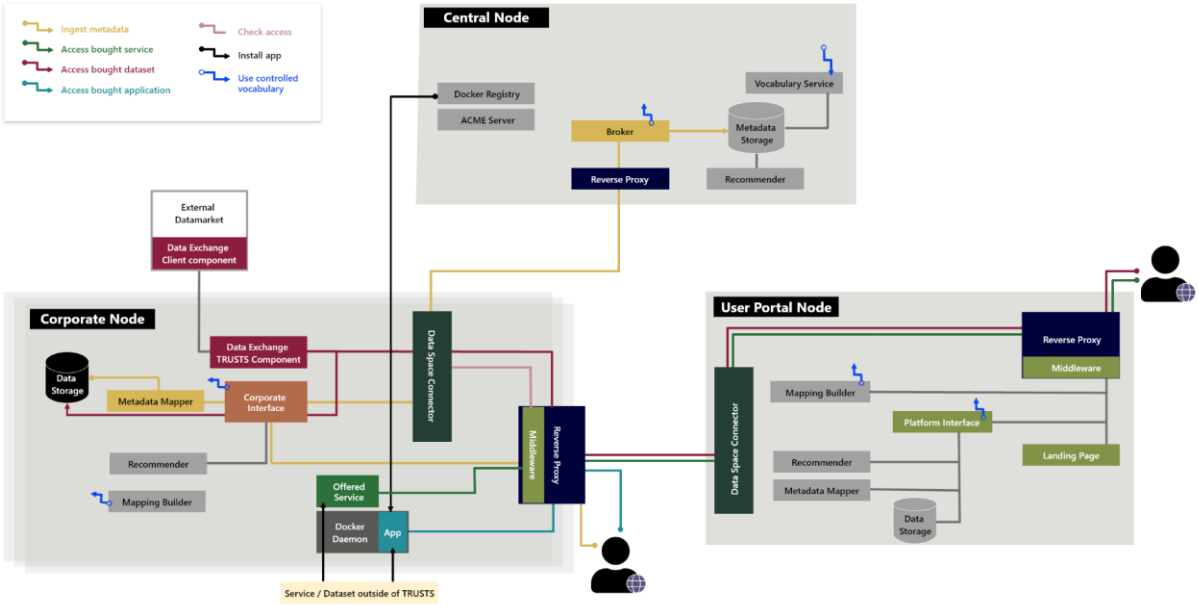

Figure 12: Overview of the architecture of the TRUSTS platform

Technical architecture of the TRUSTS Platform: Based on the architectural requirements, we describe a software architecture. We specify which software components are needed in order to address the collected functional and architectural requirements. Our proposed architecture contains 24 components. We provide tables which show that all FR and all architectural requirements are addressed by the components.

Design considerations for the architecture of the TRUSTS Platform: We describe the design considerations of the architecture to document the aspects of the architecture which are represented in the interplay of multiple components, instead of being implemented in a single component. These aspects include: the enabling of trust between participants using the platform; the functionality related to the use case trials; the handling of both portable applications and services on the TRUSTS Platform; and the planning for the evolution of the architecture with agile methods.

3.2.3 Next Steps

WP2 is completed in M24.

© TRUSTS, 2021

3.3 WP3 TRUSTS Platform implementation

3.3.1 Objectives

This WP implements the requirements and specifications for the TRUSTS platform. To achieve this, it is composed of supportive, innovative and integrative tasks. The overall objectives of WP3 are defined in the DoA as follows:

- Provision of infrastructure and operations tools and methods: Establish the technical foundations to deploy and operate the TRUSTS platform.
- Smart Contracts: Ensure the technical implementation of a smart contract feature in compliance with according regulations.
- Interoperability solutions: Provide implementations of concepts to achieve data exchange across various data market platforms and with the European Open Science Cloud (EOSC).
- Governance & Metadata: Define semantic descriptions and data models to support data interoperability, quality, lineage and data governance.
- Integration: Compile the results into a deployable, unified TRUSTS platform solution
- Brokering services: Develop intelligent recommendation algorithms that incorporate data analysis results (with respect to, e.g., platform interactions, or service description) in order to find and suggest potential collaboration opportunities between parties.

This WP aims to address the relevant requirements identified in WP2 by providing implementations in the form of software artifacts, metadata artifacts and documentation. In particular, the FR from task T2.2 and the requirements for testing and benchmarking from T2.3 guide the work in this WP. WP3 is also in close alignment with WP4, where privacy preserving technologies are investigated and developed to enable TRUSTS to provide a safe, private and trustworthy environment for the UCs with high requirements in the areas of security and privacy.

The results of WP3 are used in WP5 in order to provide the foundation for the UC trials. Three different UCs based on real world scenarios and involving a realistic subset of relevant stakeholders are set to be used to demonstrate the potential for TRUSTS.

During the second year of the project, WP3 focused on:

- Improving the development environment for the platform using state-of-the-art cloud infrastructure.
- Developing a demonstrator for smart contracts.
- Starting the development of the components responsible for data interoperability as part of the TRUSTS platform.
- Starting the development of the components related to data governance and recommendation as part of the TRUSTS platform.
- Testing and setting the second version of the minimum viable prototype for the TRUSTS platform and testing it together with the use case partners.

3.3.2 Progress achieved

Task 3.1 Infrastructure set-up and technical operations

The aim of this task is to provide a quick start environment for the development of the TRUSTS platform components. The platform provides a set of capabilities that enable operators to develop applications with a high degree of privacy-by-design features. The cloud-based environment for the development infrastructure set-up is using Google Cloud, which is compliant with the European laws and offers robust servers with tools

© TRUSTS, 2021

to ensure data security with backup, monitoring and encryption. All the resources used in the project are located in Google's EU servers. If necessary, we will liaise with the legal partners to clarify the compliance between the legislation in EU and US, should this be required.

LSTech is employing DevOps and all the state-of-the-art mechanisms to support the architecture and specifications defined in T2.4 "Architecture design and technical specifications". The Docker environment allows easy implementation, extensibility, portability, and security, allowing the different containers to run quickly from one computing environment to another.

In the second year of the project, T3.1 continuously seeked feedback from the project partners with a technical perspective on the platform, in order to improve and iterate on the provision of infrastructure. In particular, this was based on the experiences of the tasks in WP3, WP4 and WP5. The resulting second iteration of the infrastructure has been documented in deliverable D3.2 "TRUSTS Infrastructure II" submitted in M24.

More specifically, actions and tasks that were performed during the previous year include:

- Setup and manage the development environment,
- Supporting the needs of the partners in terms of resources, roles, access, deployment instructions and documentation,
- And hands on support, using CI/CD pipelines (using Jenkins) and gitlab repository to ease the development.

Moreover, ensure the security, availability and integrity of the platform through limited/ controlled access, role management, versioning and backup and restore procedures.

Finally, T3.1 provides an issue tracking and management system, an instance of the redmine software to assist the monitoring and development of the related issues.

Task 3.2 Smart Contracts

Concrete UCs for the use of smart contracts within the TRUSTS ecosystem were created using the architectural design of D2.6 analyzing how smart contracts could be used in TRUSTS.

In addition, T3.2 established some architecture requirements, including the need for certain policy definitions between a data consumer and a data provider, on the basis of which this task defines the smart contracts. The first draft of these policy definitions contains:

- Provide Access
- Inhibit Access
- N Times Usage
- Duration Usage
- Usage During Interval
- Usage Until Deletion
- Perpetual Access (Payment once)
- Access Rental (Recurring Payment)

Based on this, T3.2 was able to get a better picture of how a transaction takes place there and how it could be supported by smart contracts. The insights from this were transferred into the deliverable besides general information regarding technical, security and legal aspects of smart contracts. In addition, T3.2 generated draft contracts, on the basis of which smart contracts will be developed later on.

© TRUSTS, 2021

In the next step, this task started to set up a blockchain demonstrator. For this, existing work and frameworks for smart contracts were evaluated in order to exploit synergies within the ecosystem. To manage all the requirements, T3.2 decided to use Hyperledger Fabric (HLF) as blockchain technology for the demonstrator. HLF has some advantages: It is open source, has extensive documentation and also supports smart contracts, which is the most important point. Furthermore HLF is an enterprise-level distributed ledger and blockchain. It is designed to be interoperable with a variety of auxiliary services such as privacy, consensus, certification authority and membership service providers. T3.2 also integrated the Hyperledger Explorer UI which displays metrics about the operations on the blockchain, as seen in the below figure.

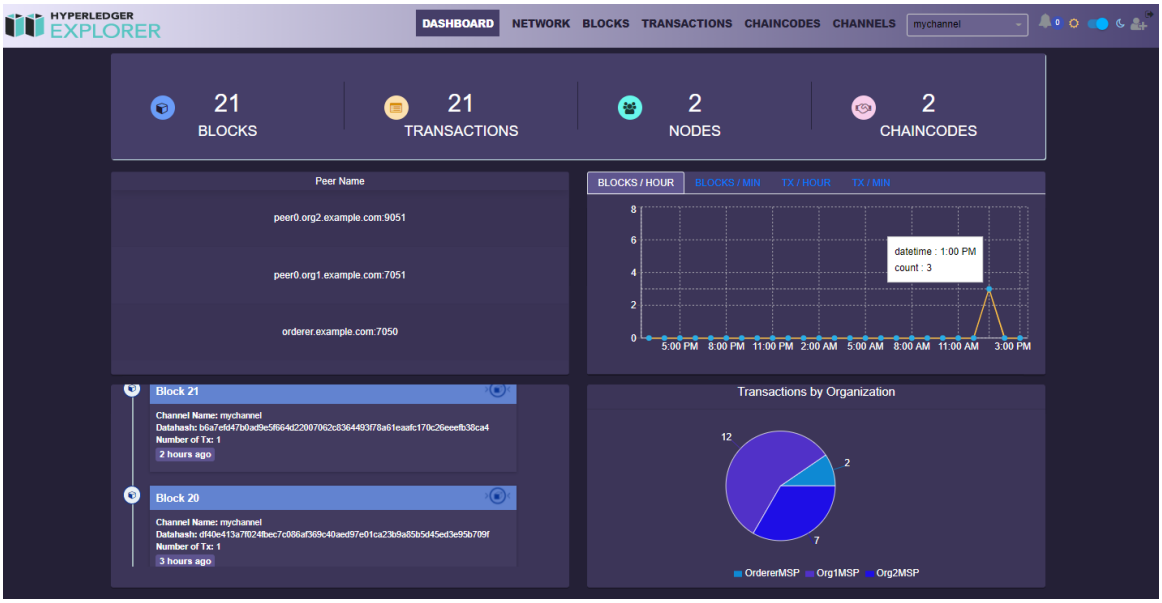

Figure 13: Hyperledger Explorer UI

The smart contract demonstrator being developed within this task consists of a blockchain instance, a blockchain monitoring UI and a library of example smart contracts which, when complete, will carry out core functions necessary to the TRUSTS Platform. It is currently possible to execute asset transfer and blockchain query smart contracts. The development plan for the demonstrator includes compatibility with a payment system via API as well as adding more smart contracts to cater for essential TRUSTS Platform operations defined by liaising with consortium partners. Integrating the demonstrator into the wider TRUSTS Platform is a target development milestone and the intermediary step towards this is the development of a client application which can connect remotely to the demonstrator using one of its exposed APIs and use the smart contract functionality. A high-level overview of the current demonstrator architecture is included in the below figure 13.

© TRUSTS, 2021

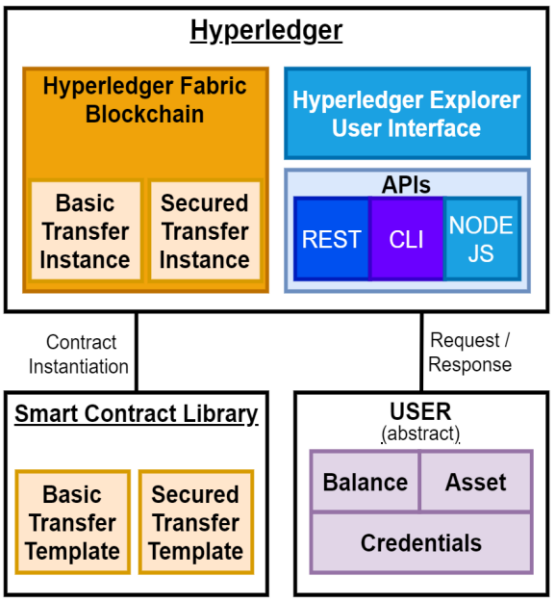

Figure 14: Current Demonstrator Architecture

The following figure gives a rough overview of how the communication between the demonstrator and a company is performed. At the beginning a user and a company conclude on a contract about a service or a dataset. After this is done, a smart contract within the demonstrator will be triggered by the company, passing all the necessary information of the contract. Each time the user now wants to retrieve the service or dataset of the company, the company is able to verify whether the user is authorized to use it or not.

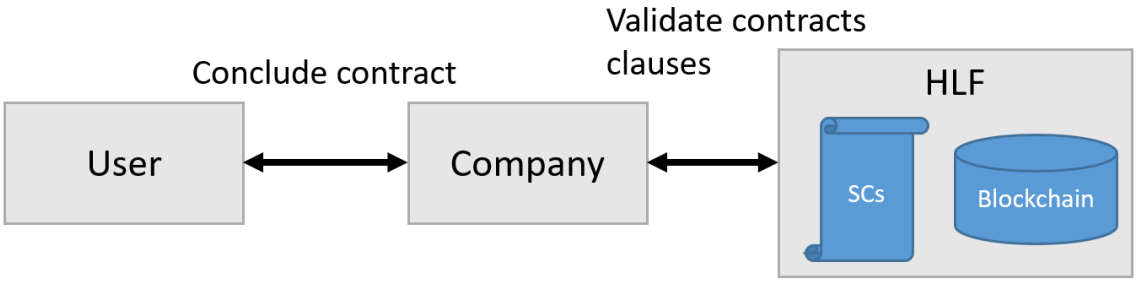

Figure 15: Rough Visualisation of the Communication

In addition to the demonstrator, D3.3 was continued. The focus layed on the general part with background information of the different technologies like blockchain and smart contract.

Task 3.3 Data marketplace interoperability solutions

The interoperability solution envisaged in T3.3 aims to interoperate with external data markets on the one hand and EOSC initiatives on the other hand. In the following, we describe our attempts to better understand the technical requirements of data markets and our work with regards to EOSC.

Systematic review of data management platforms

The systematic review of data management platforms helped to better understand the technical basis of potential datamarkets. Based on the experience of a previous deliverable (D3.4 Data Marketplaces with

Interoperability Solutions I), it is considerably hard to learn and understand the technology stack of existing platforms. Therefore, we decided to accomplish a review of libraries (software tools) that can theoretically serve as the technical backbone to build such a platform. For example, TRUSTS itself uses the data management platform CKAN as a backbone, which was also part of this systematic review.

The review followed the scientifically accepted approach of design science. According to this approach, we divided the process of review into the segments “Problem identification & motivation”, “Objectives of a solution”, “Design & development”, “Demonstration”, “Evaluation”, and “Communication” (see Figure 15 for an overview of this process and interdependency between the different segments).

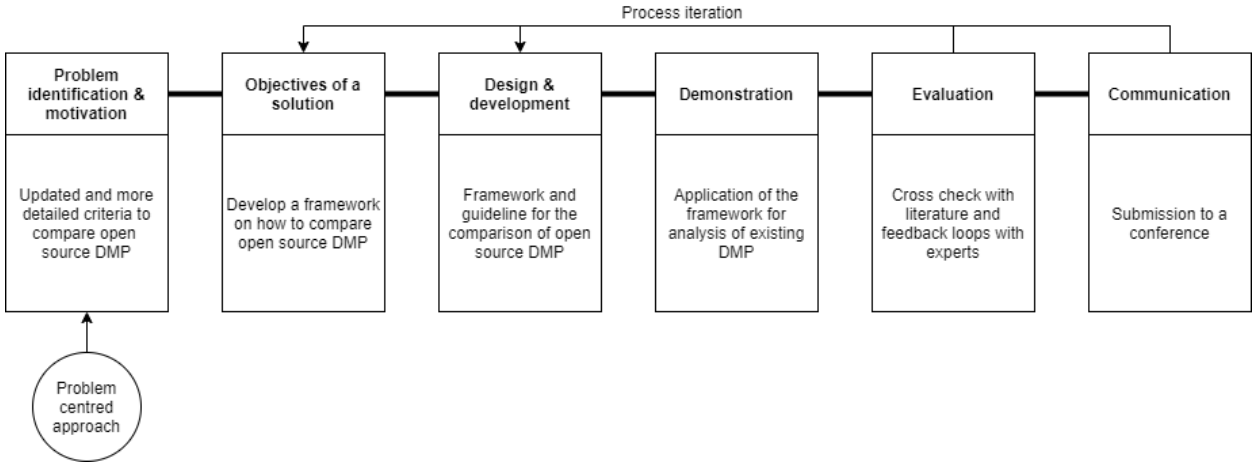

Figure 16: Overview of the design science process

Based on these segments, T3.3 derived the five iterations displayed in Figure 15. The first step was a literature review to find out about existing systematic reviews and learn about technical details of such platforms. This research showed a significant lack of existing material and the requirements given in TRUSTS, and provided the legitimacy to further conduct scientific work in this direction. Based on existing literature, we extracted a first criteria catalog to compare existing platforms. In the second step, a group of experts (from RSA), reviewed this first criteria catalog. Suggestions and comments were included in the criteria catalog. Subsequently, the criteria catalog was applied to a first, limited set of platforms. With both the feedback from experts and the practical application the final criteria catalog was established. The final step 5 then consisted of the application of the finalized criteria catalog onto a selected set of platforms. Two experts assessed each of the platforms. This double-assessment provided insight into the robustness, contrary classifications were examined by a third expert, who had not classified the respective platform so far.

© TRUSTS, 2021

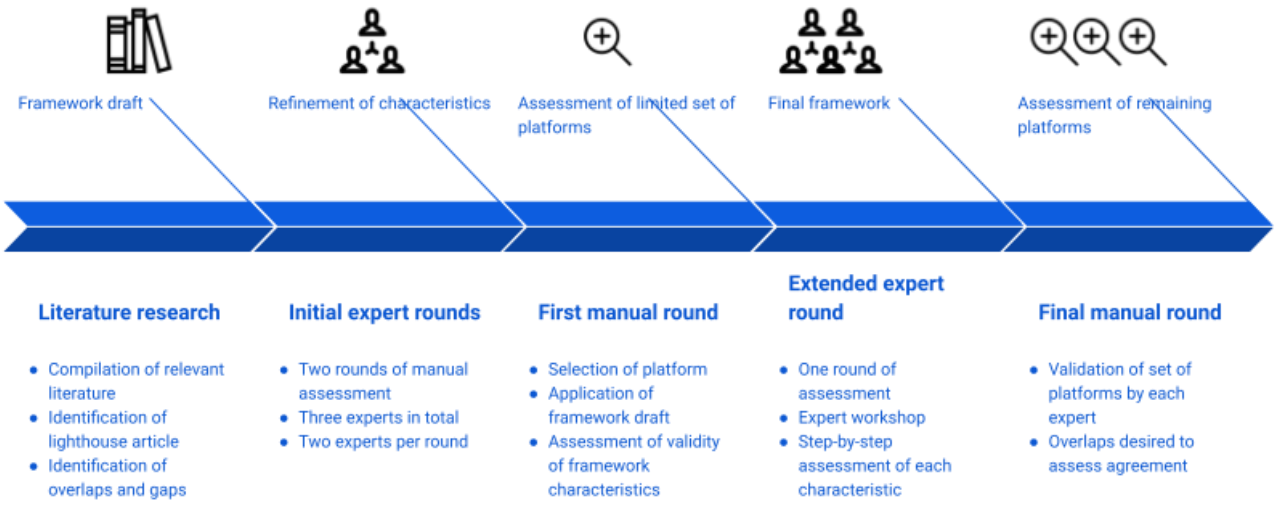

Figure 17: The individual steps of the DMP assessment.

The result of the described process was (i) a criteria catalog, applicable for further research in the area of data management, and (ii) the classification of data management platforms. The finished criteria catalog has 40 items, i.e. criteria, in total. We applied it to the seven data management platforms, i.e. we assessed the platforms CKAN[1] (which is the backbone of TRUSTS), Dataverse[2], DSpace[3], ePrints[4], Fedora[5], InvenioRDM[6], and Omeka[7]. Table 4 shows an excerpt of this assessment. The full assessment is provided in TRUSTS "D3.5 Data Marketplaces with Interoperability Solutions II". We also submitted a publication at World CIST 2022[8] covering the criteria catalog, the process of creation, and the finished assessment.

Table 4: The first ten results of the systematic review.

	Feature	CKAN	Dataverse	DSpace	ePrints	Fedora	InvenioRDM	Omeka
1	Installation wizard	✖	✓	✖	✖	✓	✖	✖
2	Installable from repository	✓	✓	✓	✓	✓	✓	✓
3	Container	✓	✓	✖	✓	✓	✓	✓
6	License	AGPLv3	Apache V2	BSD	LGPLv3	Apache.2.0	MIT	GPLv3
7	Ecosystem of extensions	✓	✓	✖	✓	✖	✖	✓
8	Internationalization support	✓	✓	✓	?	✖	✖	✓
9	Multi factor authentication	✓	✖	✓	✖	✓	✖	✖
10	Authorization (access control)	✓	✓	✓	✓	✓	✓	✓

[1] https://ckan.org/, last accessed Nov 26, 2021.
[2] https://dataverse.org/, last accessed Nov 26, 2021.
[3] https://www.dspace.com/de/gmb/home.cfm, last accessed Nov 26, 2021.
[4] https://www.eprints.org/uk, last accessed Nov 26, 2021.
[5] https://duraspace.org/fedora/, last accessed Nov 26, 2021.
[6] https://inveniosoftware.org/products/rdm/, last accessed Nov 26, 2021.
[7] https://omeka.org/, last accessed Nov 26, 2021.
[8] http://worldcist.org/, last accessed Nov 26, 2021.

© TRUSTS, 2021

EOSC

Interoperability with EOSC is the second major aspect of T3.3. EOSC, the European Open Science Cloud[9], is not a single platform or data market for science-related data sets. Instead, it is an umbrella for a plethora of science-related initiatives. EOSC sorts its initiatives along the eight domains "Medical & health sciences", "Engineering & technology", "Natural sciences", "Generic", "Humanities", "Agricultural sciences", "Social sciences", and "Other". Additionally, the initiatives are also sorted along the categories "Access physical & eInfrastructures", "Aggregators & integrators", "Process & analysis", "Security & operations", "Sharing & discovery", and "Training & support". The initiatives do not only differ in their goals and scientific orientation, but also with regards to their technological background. They do not share a common interface, which TRUSTS could leverage. Instead, they are different web applications for different purposes, e.g. search engines, provision of virtual machines, or cloud-based research. We have conducted intensive research to get an overview of the diversity of EOSC initiatives, and provide an exemplary overview in Table 5 for the domains "Engineering & technology" and "Medical & health sciences". The full results are available in TRUSTS "D3.5 Data Marketplaces with Interoperability Solutions II".

Table 5: Examples of the technical background analysis of a set of EOSC initiatives.

Initiative name	Short description	Type of resource	API?
Engineering & technology			
Europeana APIs	Large-Scale Data Discovery, Acquisition and Management of Digital Cultural Heritage in Research	Data Retrieval API Set (Free – registration required)	Y – several (https://pro.europeana.eu/page/apis#our-apis)
MetaCentrum Cloud	Czech national scientific cloud	IAAS for scientific users (Free – registration required)	Y (https://cloud.gitlab-pages.ics.muni.cz/documentation/register/?q=API)
Medical & health sciences			
3DBIONOTES-WS		Web application	Y (http://3dbionotes.cnb.csic.es/ws/api)
AMBER-based Portal Server for NMR structures (AMPS-NMR)	Web portal for the refinement of Nuclear Magnetic Resonance (NMR) structures of macromolecules	Web portal	No

T3.3 will tackle the diversity in technologies in two ways: on the one hand, we will leverage the harvesting functionality of CKAN. CKAN is the backbone of TRUSTS, i.e. it can use a CKAN extension for harvesting. Multiple EOSC initiatives are built with CKAN (e.g. B2FIND[10] and the EOSC pillar catalog[11]), which means that both TRUSTS and these initiatives can interoperate with each other (minor modifications required). On the

[9] https://eosc-portal.eu/, last accessed Nov 26, 2021.
[10] http://b2find.eudat.eu/, last accessed Nov 26, 2021.
[11] https://ckan-eoscpillar.d4science.org/, last accessed Nov 26, 2021.

© TRUSTS, 2021

other hand, we will attempt to become a so-called "EOSC provider"[12]. This requires a registration process as well as the adherence to the technical requirements given by EOSC. Upon completion, TRUSTS can become an active provider of data assets to EOSC.

Task 3.4 Data Governance: Metadata, Lineage and Semantic Layer

Task 3.4 aims at the definition of data governance and metadata management practices, as well as their implementation. After an analysis of the different requirements, and the available standards and approaches to satisfy them, task T3.4 has compiled a proposal of the TRUSTS Information Model. This proposal is based on the work done in the IDS project, with adaptations specific to the technical and business requirements of the TRUSTS Platform. The result, as documented in deliverable D3.7, is an ontology which specifies the way in which different assets in the TRUSTS Platform are described, including guidelines on how these descriptions are to be interpreted by the different components. The collections of these descriptions are known as the TRUSTS Knowledge Graph, to be stored in the Metadata Storage component.

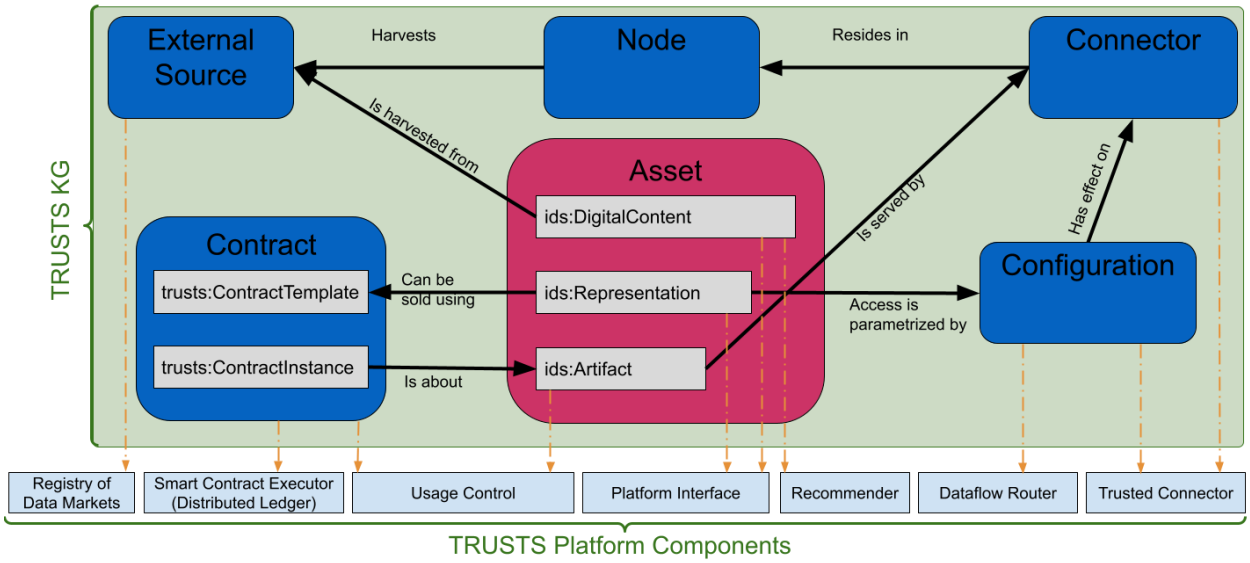

Figure 18: Overview of the TRUSTS Knowledge Graph

Great attention has been put into the description of assets as required by the federated nature of the TRUSTS Platform, the forwarding of requests between different participant nodes, and the interaction with external third-party marketplaces.

[12] https://eosc-portal.eu/for-providers, last accessed Nov 26, 2021.

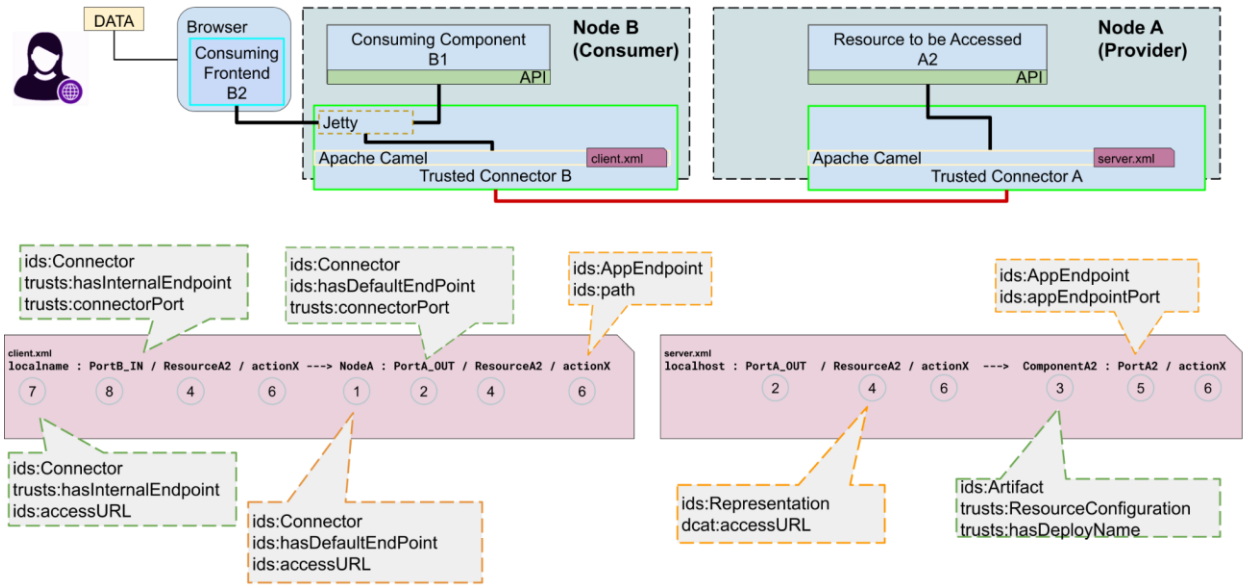

Figure 19: Overview of asset descriptions required by the federated nature of the TRUSTS Platform

Alongside the definition of the TRUSTS Information Model, a study and selection of the appropriate controlled vocabularies was undertaken. These vocabularies are meant to be used as possible values in several of the fields specified in the Information Model so that, for example, the format of a file or the type of access method is one from a list of options, each of which has well-defined semantics and can be consistently interpreted by any involved components.

The above mentioned developments have been coupled with adaptations and deployments of several software components, both during the definition and the test phase of the Information Model and the corresponding vocabularies. In this direction, an extension CKAN, a popular data-portal software, has been developed to make it compatible with the TRUSTS Information Model. Furthermore, configuration of the IDS Trusted Connector was done in accordance with the model (an exercise which will be useful for the configuration of routing mechanisms in other connectors or similar solutions), and test deployment of this Connector, coupled with CKAN has been done.

Finally, architectural discussions (T3.5), as well as those concerning interoperability with external data sources (T3.3), were constantly informed with the metadata perspective. This has resulted in contributions to deliverable D2.6, as well as organization of meetings among different tasks and work packages in order to standardize nomenclature and understanding of the different functionalities.

Task 3.5 Platform Development & Integration

Task 3.5 continues the platform development and integration during the second year of the project. During this period the next version of the minimum viable prototype (MVPv1) was developed based on the IDS Trusted Connector and the CKAN framework. This infrastructure was developed to support secured communication between nodes and nodes to the central node of the platform. Below you can see the communication diagram of TRUSTS platform.

© TRUSTS, 2021

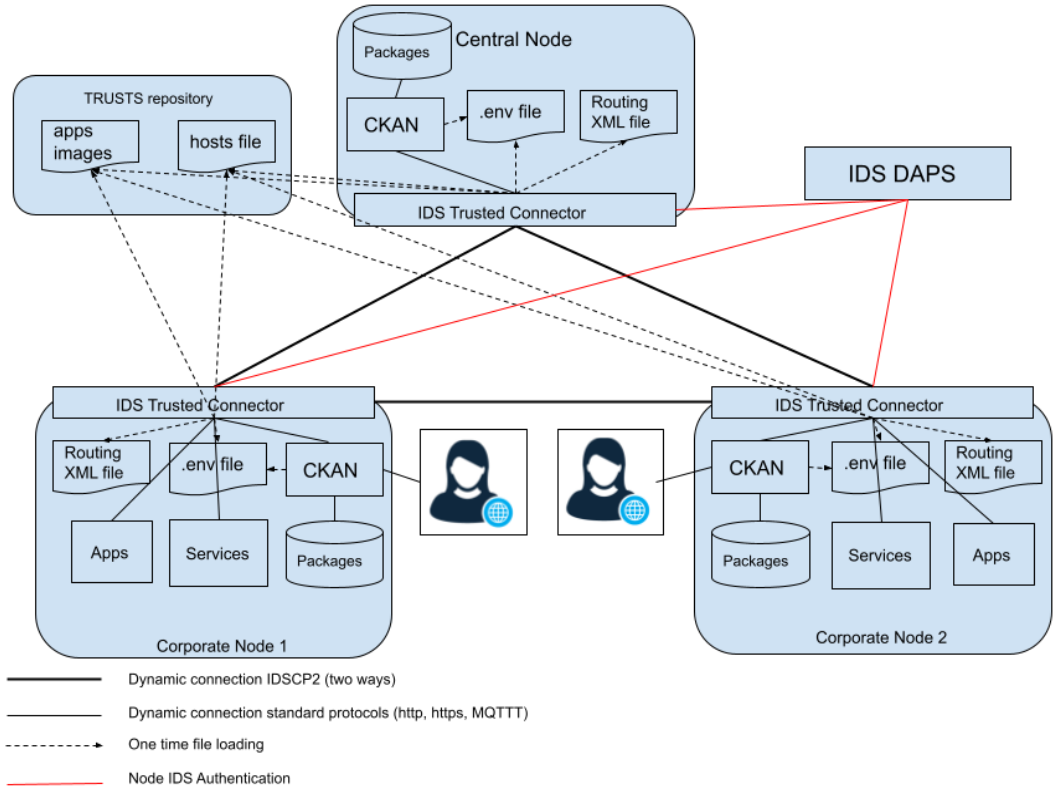

Figure 20: Schematic architecture of minimum viable prototype version 1 (MVP.v1)

MVPv1 was deployed in gcloud TRUSTS infrastructure, tested and the UC partners were trained how to use it. MVPv1 supports providers for creation of datasets, services and applications with needed supporting information. In addition, it makes assets available for searching and usage. A consumer can download datasets and packages for services and applications and can receive secured access to services through TRUSTS platform.

The second half of the year was spent on improving the platform. In addition, new open source releases from the IDSA were incorporated into the ongoing development of the MVP version of the platform. These components are the IDSA Dataspace Connector and the IDSA Metadata Broker. Additionally to them T3.5 added open source reverse proxy Traefik and open source Certification Authority Small Step CA.

Task 3.6 User and corporate profiles and brokerage

T3.6. aims to design and set up brokerage services in the form of a recommender system for interlinking user and corporate profiles with services (including applications) and datasets available from within the TRUSTS platform. Based on the FR of the TRUSTS platform, we have further refined the recommendation use cases for fulfilling these requirements. This leads to the final set of 6 recommendation use cases:

1. RUC1: the recommendation of datasets to users,
2. RUC2: the recommendation of services to users,

© TRUSTS, 2021

3. RUC3: the recommendation of datasets to services,
4. RUC4: the recommendation of services to datasets,
5. RUC5: the recommendation of datasets to datasets, and
6. RUC6: the recommendation of services to services.

Additionally, T3.6 set up the technical infrastructure to implement these recommendation UCs, which is depicted in the following figure and which is described in detail in deliverable D3.12 that was submitted in M18:

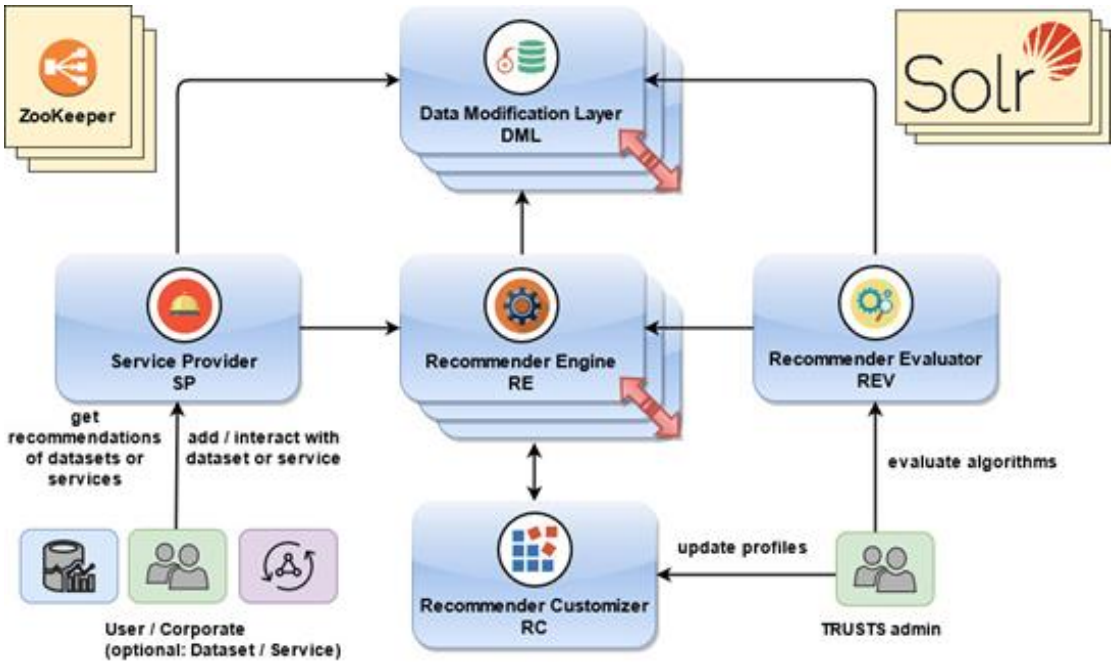

Figure 21: Overview of technical infrastructure for recommendations as part of the TRUSTS platform

In close collaboration with Tasks T3.4. and T3.5, T3.6 ensured that this technical infrastructure is able to consume data generated in the TRUSTS platform in line with the IDS information model. This includes that the recommender system interacts with the metadata catalogue of the IDS broker (for receiving interaction data and metadata for users/services/datasets) as well as with the TRUSTS portal (for showing recommendations and receiving feedback of recommendation clicks).

As also described in D3.12, we have conducted research in the area of privacy-aware recommender systems. Here, we aim to provide accurate recommendations with a reduced amount of private user data. This work was published and presented in the reproducibility track of the European Conference on Information Retrieval (online conference): Muellner, P., Kowald, D., and Lex, E. (2021). Robustness of Meta Matrix Factorization Against Strict Privacy Constraints. In Proceedings of the 43rd European Conference on Information Retrieval (ECIR 2021). Springer (https://arxiv.org/abs/2101.06927). Additionally, T3.6 presented this work at the Responsible AI Forum in Munich (Germany).

3.3.3 Next Steps

Task 3.1 Infrastructure set-up and technical operations

Next steps include the deployment of Google Kubernetes Engine in order to guarantee robustness into the deployment stage. As the project is growing, cluster orchestration software will be implemented on a layer

above container management (Kubernetes). This will ease the deployment of the whole infrastructure as one. Also benchmark components and load balancers will be set up. Ansible and Puppet are the suggested technologies.

Task 3.2 Smart Contracts

The next steps are to finish the demonstrator and make the results and findings available to the partners on the one hand and to integrate them into the deliverable D3.3 on the other hand. Furthermore the writing of D3.3 will be continued to complete the concept of using smart contracts in TRUSTS in time.

Task 3.3 Data marketplace interoperability solutions

The next steps cover the implementation of the envisaged components (one component, the registry of data markets, is already finished). This includes the development of the Data Exchange TRUSTS component and the Data Exchange client component. Furthermore, implement the interface that allows external entities to connect with TRUSTS and share their data catalogs with TRUSTS.

Task 3.4 Data Governance: Metadata, Lineage and Semantic Layer

The next steps are centred around the implementation of the metadata ingestion platform, which will implement the following functionality:

1. prototypical ingestion of metadata into a Broker
2. pipeline for the ingestion of metadata related to services and datasets, and
3. provision of metadata endpoints for the management of other TRUSTS assets such as nodes and organizations.

This implementation entails the setting up of specialized computing infrastructure, the selection of use cases, and the establishment of acceptance criteria both for individual components and for integrations.

Task 3.5 Platform Development & Integration

T3.5 will be further iterating the TRUSTS platform. This will be based on interactions and feedback with the other tasks in WP3, and on interactions and feedback with the UCs in WP5. Another aspect will be the definition and testing of security configurations to provide the necessary APIs to enable WP4 to implement their requirements. This also involves setting up a testing environment with instances of all components required for providing a secure data marketplace and secure services. Finally, T3.5 will also facilitate the creation of a development and testing environment for the services hosted on the platform with the help of the general infrastructure provided by T3.1.

Task 3.6 User and corporate profiles and brokerage

In the third and final year of TRUSTS, T3.6 plan to work on three concrete tasks:

Firstly, and from a technical point of view, the aim is to fully integrate the recommender system into the IDS-based infrastructure of TRUSTS. Specifically, this will lead to the implementation of the 15 services shown in the following figure and described in deliverable D3.12:

© TRUSTS, 2021

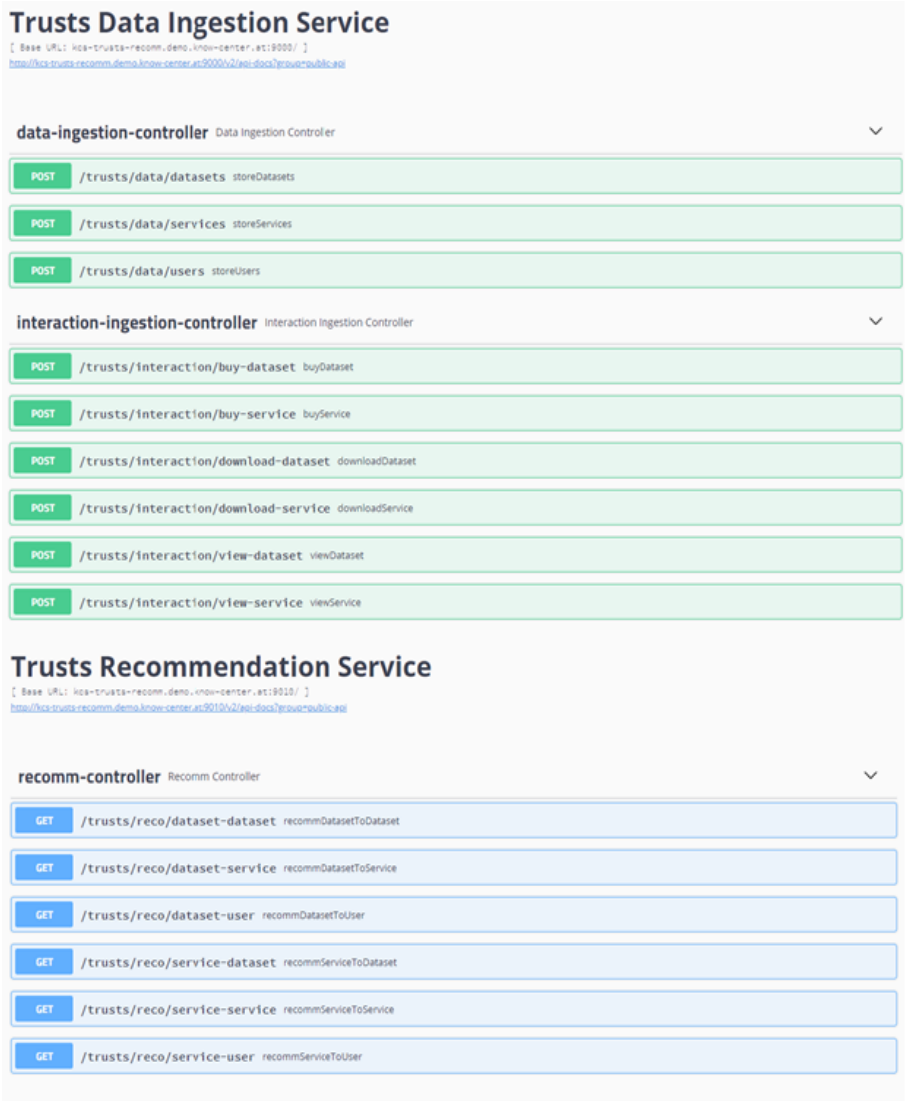

Figure 22: Mockup user interface of the services provided by the recommender component

Secondly, and from an evaluation perspective, T3.6 will use publicly available data from the OpenML Machine Learning (ML) platform (https://www.openml.org/) to evaluate the recommender system. This will also allow the fine-tuning of the algorithms. Thirdly, and from a research perspective, T3.6will continue the research on privacy aspects of the recommender system. This should lead to the goal to enhance the algorithms implemented in the recommender system with respect to the privacy-accuracy trade-off. All these developments will be described in the concluding deliverable D3.13, which is due to M36.

© TRUSTS, 2021

3.4 WP4 Privacy preserving technologies

3.4.1 Objectives

Data privacy is a worldwide concern due to threats and risks that can compromise individual security, reputation, and social exposure. Aiming to mitigate privacy risks and assure civil rights on personal data, countries and territories have been ruling the activities related to data collection, transfer, storage, management, and deletion, such as the European Union's General Data Protection Regulation (GDPR). The terms of such regulations also play a crucial role in the development of artificial intelligence models that explore personal data, namely Machine Learning (ML) and Deep Learning (DL). Therefore, mechanisms for privacy preservation such as differential privacy, homomorphic encryption (HE), Federated Learning (FL), secure components, multi-party computation (MPC), and adversarial training have been successfully proposed and applied to real-world systems. Furthermore, tools for measuring data de-anonymization and/or mitigating this anonymizability have been proposed to tackle the problem.

Advanced decision-making capabilities are required for broad areas and arise from the improvements of data science along with the technical ability to draw advanced conclusions based on big data. These capabilities have been proven when it comes to public data. However, for private or personal data, there still exists the requirement to develop a technology platform that will allow the execution of advanced techniques for data analytics alongside the complete prevention of data breaches that may endanger privacy. Furthermore, regulatory constraints and the desire to preserve individuals' privacy uphold the accomplishment of this requirement, which is the main objective of this project.

This WP has the objective of integrating privacy-preserving mechanisms to TRUSTS in order to safeguard the UCs in the financial domain from privacy threats. In addition, data trading and sharing activities will also be protected. Furthermore, WP4 has the objective to provide tools for anonymization and de-anonymization.

Because personal private data trading is not possible in the ordinary sense of the word, WP4 is required to develop the ability to support data processing without compromising data privacy. Throughout the project WP4 works in full collaboration with the UC leaders and the WPs leaders in order to adapt the research perfectly to the system requirements and the UCs requirements.

3.4.2 Progress achieved

T4.1 and T4.2 were fully finalized and the outcomes of it were reported in D4.1. [M18]

T4.1 Privacy Preserving Data Analytics

In contrast, MPC is a technology that allows computations with two or more input parties. This much greater flexibility comes with the cost that each party has to take part in the computation actively.
To sum up, for a client-server, setting FHE is a proper choice. In all other cases, one would use MPC-protocols, sometimes also combining them with FHE.

Cryptographic primitives involved in building collaborative trust systems were investigated. ***Fully homomorphic encryption (FHE) –*** Setting up an FHE framework will allow you to do outsource computation without giving up any privacy and without having to trust the service provider, since they are not able to access the actual content of your data.

Secure multi-party computation (MPC) – SecureMPC provides similar confidentiality and privacy in the real-world, where one cannot fully trust third parties. Therefore, what can be achieved in the ideal-world, can also be done by applying secure MPC.

Private Set Intersection – Private set intersection is a special-purpose secure MPC. It allows two participants to compute the intersection of their data sets.A PSI application was developed.

Homomorphic Encryption versus Multi-Party Computation – FHE is an appropriate choice in the classical client-server setting, whereas MPC prevails whenever at least two parties actively perform a computation.

The nature of FHE fits an infrastructure that can be widely found on the internet. On the one side, we have a weak client device, like an edge device, a smartphone, or a laptop with input data. On the other side, there is a powerful server (computationally or application-wise). Usually, the client provides input data to the server because it offers a useful application to the client (which the client cannot do on his own). In this scenario, the server performs all the computation. The client only provides input data and could go offline during the computation. This perfectly matches the FHE design. In FHE, the client encrypts its data, sends it to the server, which does all the computations, and then sends back the result to the client.

T4.3 Anonymization and de-anonymization

The objective of Task 4.3 is to provide tools for risk analysis and anonymization that help in safely possessing personal data. In particular the aim of the task is to help in the following:

1. Raising awareness: Laypeople that do not have prior knowledge on anonymisation view it as removing the PIIs only and are not aware of the de-anonymisation risks in their datasets. De-anonymisation risk analysis tools can reveal and highlight the de-anonymization risks in datasets and the extent to which they are de-anonymisable.
2. Compliance to GDPR: As described previously, GDPR does not specify which privacy models are suitable in which cases and implies that the data controller should become aware of the de-anonymisation risks in their datasets. De-anonymisation risk analysis tools help in the compliance to GDPR since they report on how much datasets conform to privacy models and raise the awareness of the de-anonymization risks.
3. Aiding in the anonymisation measures and their extent: Deciding the anonymisation measures and their extent is the core challenge in the anonymisation process. De-anonymisation risk analysis tools, if designed properly, can help in this decision and its extent since they can reveal the distortion that is required for a dataset to comply with privacy models. Examples are described in the next section.

T4.4 Federated Deep Learning methodologies [M18-36]

This task constitutes a horizontal layer of the TRUSTS architecture facilitating the federated training and utilization of the envisaged DL algorithms, which will be incorporated in the platform, by distributed devices, running on the edge of the system's cloud. A cloud based framework will be deployed enabling the distribution, training, inference, monitoring and update of existing AI models to selected distributed clients, which will be able to utilize local isolated content repositories. To this end, each federated deployment is enabled to use private or sensitive datasets for the generation of the necessary feedback to the TRUSTS

© TRUSTS, 2021

platform, without endangering their unauthorized access or exposing the data source.

T4.5 Transformation of algorithms to privacy-preserving certified [M18-36]

This task will strive to convert risky algorithms that compromise privacy into safe and privacy-preserving without harming their functionality. Various algorithms ought to use external sources and run computation to execute certain functions. The development of most algorithms is driven by outcome and performance, leaving privacy and security issues on the least of requirements. The challenge is in retrofitting and enabling working algorithms to perform under the desired set of privacy regulations without the need of redevelopment.

T4.2 Privacy Preserving Transfer Learning and Classification

Task 4.2 Privacy Preserving Transfer Learning and Classification, which has been running from M1 to M18, presented the challenge of bridging the gap between Transfer Learning (TL) and privacy-preserving methods of HE and DP for financial datasets. As a result, a private, efficient, and secure TL method, namely CryptoTL, was proposed and had its efficiency tested over publicly available benchmarks datasets for the credit risk assessment task. Beyond coming up with novelty regarding the combination of TL with HE and DP, this framework has potential to be applied to many other real-world use cases. As future works, CryptoTL is expected to be extended to data types other than financial data, e.g., textual data, in order to preserve data privacy in a larger number of applications. Finally, the research outputs of Task 4.2 were submitted as a research paper for the Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS 2021).

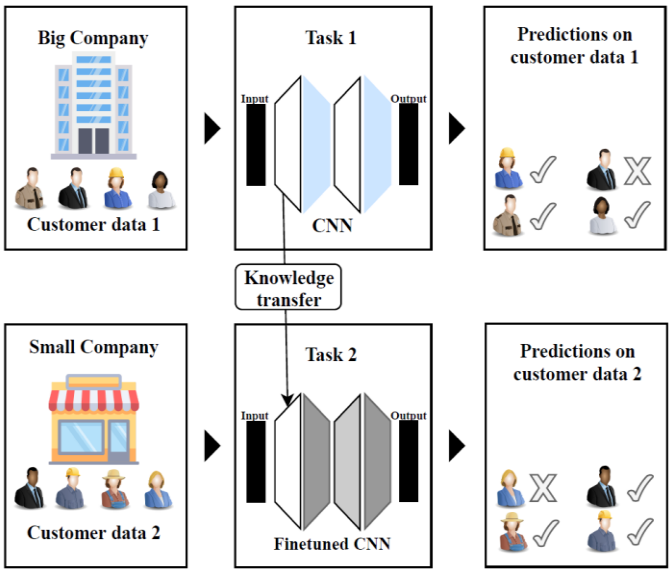

Figure 23: An Example of Knowledge Transfer between two Tasks

T4.3 Anonymization and de-anonymization

In the first year of Task 4.3, focused on developing the de-anonymisation risk analysis modules and algorithms and implementing these modules in the form of a ready-to-use application.

We build on the results of the Safe-DEED project. In the following paragraph we make clear the contribution of the TRUSTS T 4.3 from scratch and on top of the results of the Safe-DEED related tasks. Hereby, we

reference Safe-DEED deliverables D 5.6 and D 5.10. The contributions and improvements are the following:

1. Generalization: The aim in Safe-DEED was to provide ad hoc solutions to deal with particular datasets (customer relationship management (CRM) dataset provided by Forthnet). This is why the modules were subject to this data, which were seen as proof of concept. In TRUSTS 4.3, the aim by RSA was to generalize these algorithms to build ready to use modules that work with any dataset and are capable of being integrated into the application, which will be a part of the TRUSTS platform.
2. New modules: Three of the de-anonymization algorithms developed by RSA entirely in TRUSTS, namely the algorithms designed for Spatiotemporal Data and Textual Data as well as the L-Diversity algorithm, are novel algorithms that didn't exist in the Safe-DEED work.
3. Ready to use application: FORTH developed a new professional application that aims at hosting and managing datasets and the anonymization and de-anonymization modules of RSA as well as providing an interactive front end to use these modules. Such modules didn't exist in the outcomes of SafeDEED. The final goal is that this application will be a part of the TRUSTS platform

In the following, we provide a short description of each of the risk analysis modules that have been developed as well as the application that integrates them:

K-Anonymity: In this module, a tabular dataset is checked for its compliance to k-anonymity for each unique combination of its quasi identifiers (QIs). The core of this module has been developed within the EU-funded project Safe-DEED (Bampoulidis, 2020a), but it only supports k=2 and does not provide a visualisation. In this task, we have integrated and extended this module to support any k (and have improved its algorithmic complexity from O(#QIs × r) <= complexity <= O(2#QIs × r) to O(2#QIs), where r is the number of records in a dataset. Furthermore, we provide user-friendly visualization to ease interpretation of the results.

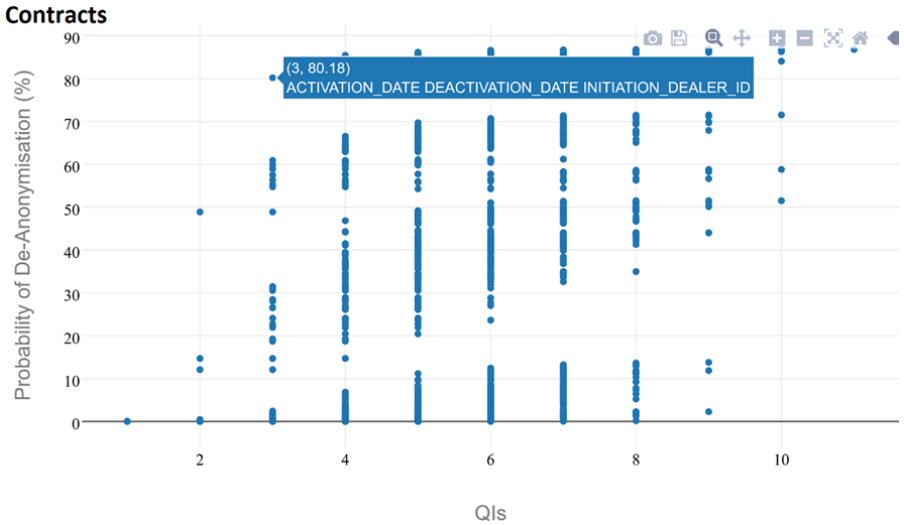

Figure 24: Probability of De-Anonymization

L-Diversity: In this module, a tabular dataset is checked for its compliance to l-diversity for each unique combination of its QIs and sensitive attributes, i.e. the probability of an adversary finding out the value of a sensitive attribute. This can lead to the same insights and actions described in the k-anonymity module. The algorithmic complexity of this module is O(2#QIs x #sensitive_attributes). L-Diversity has the same visualization technique like k-anonymity (above), but with different information.

© TRUSTS, 2021

Spatiotemporal Data: Spatiotemporal data is data that contains information about individuals' location and time. For this kind of data, we have developed the following privacy notion for the location of each individual at a specific time: w other individuals within a radius r, within a timeframe t. If an individual is located at location x at time y, and there are at least w other individuals within a radius r, within a timeframe y±t, then location x at time y is considered safe. For ease of reading, we refer to a combination of location x at time y of an individual as a check-in point. The output of this module is a map with all the individuals' locations, such as the figure below, coloured in a scale from green to red. The colour of each point represents the probability of singling out an individual (labelled as risk): how many of the check-in points at the respective location do not conform to the privacy notion specified above with w, r, t. If all check-in points at a respective location are safe (according to the privacy notion), then the location is coloured green; red, if all check-in points are unsafe.

Figure 25: An Example of Spatiotemporal Data

Textual Data: Textual data is any text data that is generated by individuals and contains personal information regarding their texting styles. The most prominent example are search logs. To measure the uniqueness of words used by individuals, we use the Jaccard Similarity to calculate the similarity of texts of two individuals, based on the Jaccard Similarity metric, which calculates a score between 0 and 1 (0 being dissimilar; 1 being similar). Doing this pairwise for all individuals results in a heat map, such as the ones below, which illustrate this analysis for logs from Amazon and AOL. The x and y axes correspond to individuals and the colour of each tile in the heatmap corresponds to the Jaccard Similarity between the texts of two individuals. Note that the Jaccard Similarity in the diagonals is 1, but the software does not render them at this zoom level.

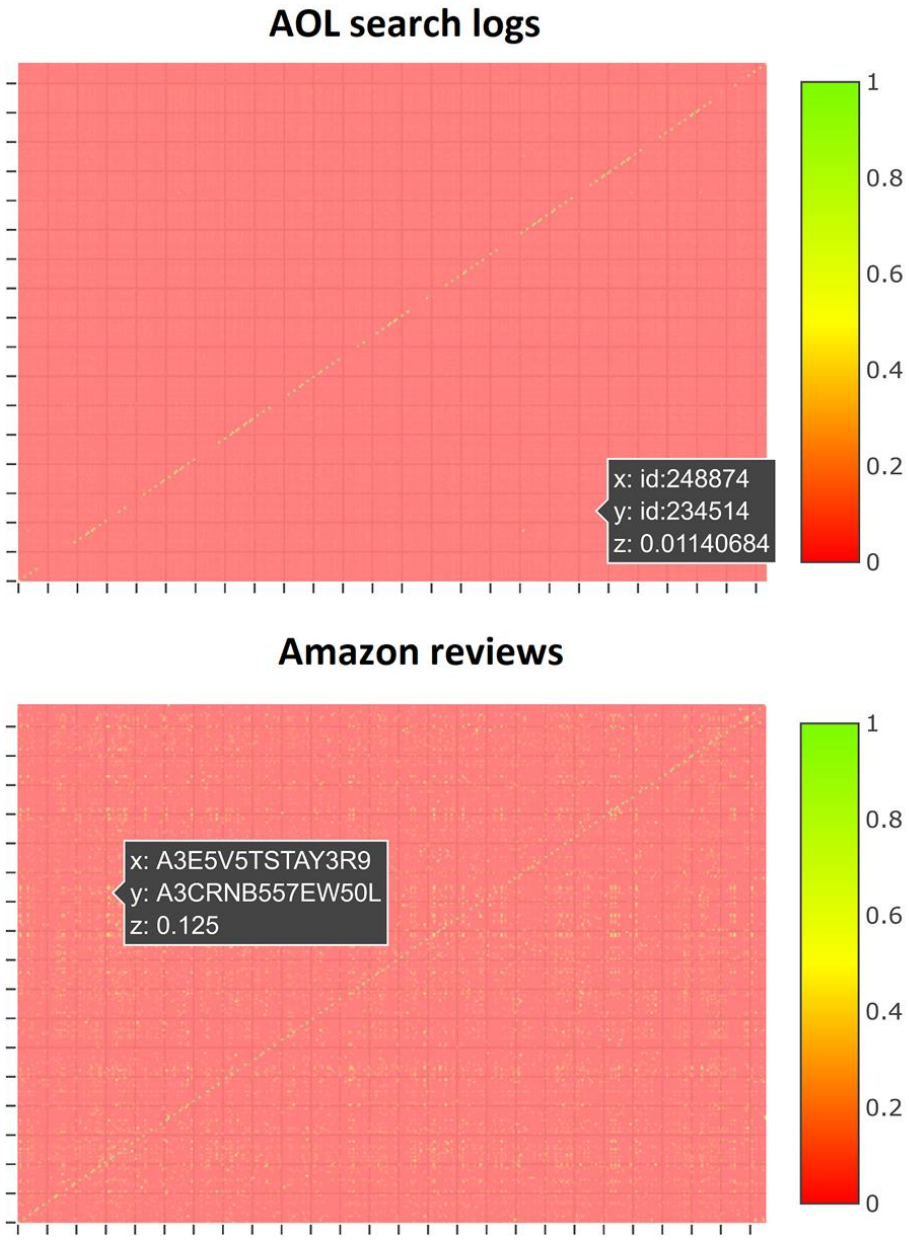

Figure 26: AOL Search Logs and Amazon Reviews

Financial Transactions Data: Financial transactions data is data that contains any information about payments individuals made or received at a specific time. For this kind of data, the following privacy notion is checked against the data: w other individuals having a transaction amount within a, within a timeframe t. If an individual has paid an amount x at time y, and there are at least w other individuals having a transaction amount x±a, within a timeframe y±t, then amount x at time y is considered safe.

The output of this module is a point plot, such as the one below, with all the unique data points of amount (x-axis) and date (y-axis) coloured green if they comply with the specified privacy notion; red, otherwise. The output of this risk analysis module reveals the outliers of the data, i.e., the individuals with distinct transactions. Additionally, the output can help the data controller decide the generalisation hierarchies (i.e.,

the binning of the amounts) in the case of anonymisation, or the aggregation levels in case of aggregation. For instance, in the figure below, there are areas with consecutive red spots that can be grouped into one or more distinct values (e.g. [100k-200k], [>200k], etc.).

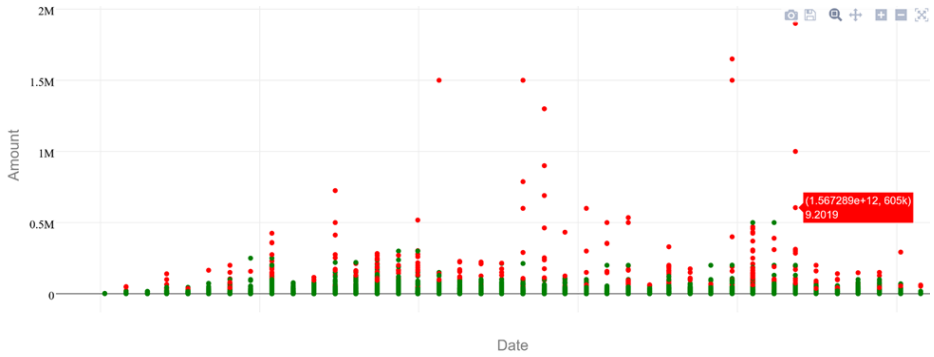

Figure 27: Point Plot for Financial Transactions

Aggregation-Based Data: Aggregation-based data is data that contains aggregate values, such as sum, count, and average, about individuals. If the aggregate values are low and there exist some other sensitive aggregate attribute, then there could be a privacy breach. The figure below is an example output of this risk analysis module: the aggregate values of a dataset visualised in a bar plot with a horizontal line (k) representing the minimum acceptable value of an aggregation attribute. In this case, the records containing the attribute "shares" being below the acceptable value should be removed, if there exists another sensitive attribute (e.g., sum of income). The core of this module has been developed in Safe-DEED and in this task we modified it to fit it to our application. The algorithmic complexity of this module is O(r), with r being the number of rows in the aggregation-based dataset.

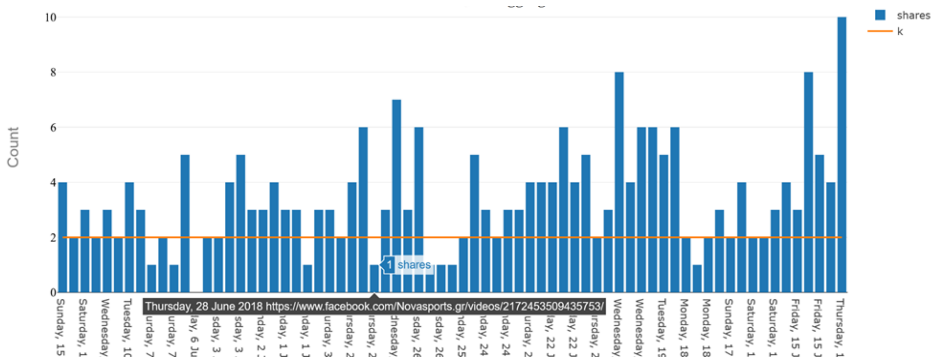

Figure 28: Bar Chart for Aggregation-Based Data

Application: The application already developed by FORTH constitutes a read-to-use toolkit, designed and developed for incorporating the above-mentioned mechanisms and providing an intuitive and usable UI to the end users. The tool has been developed as a docker container, such that TRUSTS' users can download the tool and apply risk analysis on their premises.

For architecture, the application has been developed in a loosely coupled manner, where every component is being a standalone entity in a separate docker image. All the components/images needed for the application are deployed through a single docker-compose file. The main components of the app's architecture are the following:

- The application Frontend/GUI

© TRUSTS, 2021

- The Coordinator Server, which is responsible for the necessary backend operations with regard to the coordination of the backend components according to specific workflows (e.g., data ingestion to the Data Management System from files, triggering of the risk analyses, etc.)
- The Privacy-Backend Server, which undertakes the ingestion of the dataset as well as the execution of the Risk Analysis and Anonymization processes, that are described in the previous sub-sections.
- The Data Management System between the components (metadata), storing of the resulting data of each process as well as the logging of the App's components outputs (TBI).

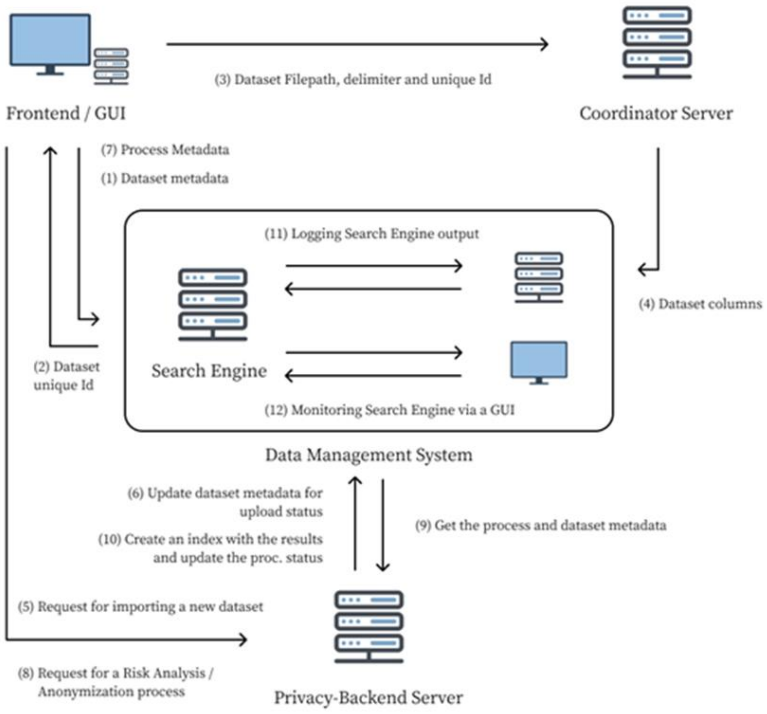

Figure 29: Main Components of the App's Architecture

T4.4 Federated Deep Learning methodologies

In general, FL can be divided into two different types: The first one is Horizontal Federated Learning (HFL) and is introduced in scenarios where data sets share the same feature space but are different in sample. This type of collaboration is very rare when it comes to different companies and different domains, but it is very common in telecommunication use cases for example.

The second is Vertical Federated Learning (VFL), which is applicable to the cases where two data sets share the same sample ID space but differ in feature space. This scenario is much more common in the industry, and it is the one that we are focusing on in the TRUSTS project.

We started with a kick-off of our research efforts focusing on VFL techniques that will enable TRUSTS parties to collaborate over their private and sensitive data while preserving data privacy. We began with a survey of recently published research papers related to privacy challenges for FL applications, such as open search and recommender systems. In these scenarios FL has been found as an efficient solution against data leakage. However, open challenges still prevail, for example related to unintended memorization of data instances by the federated model. Such data instances may represent users' personal information, preferences, or

© TRUSTS, 2021

behavior. A corresponding overview was presented at the 3rd International Open Search Symposium and we also started implementing a FL prototype for textual data.

In collaboration with Task 4.5. we also started to work on an encrypted FL version, based on a recently published multi-key HE scheme (https://arxiv.org/pdf/2104.06824.pdf). Here, the model updates of the FL system are encrypted via an aggregated public key before sharing with a server for aggregation. This greatly enhanced the security of typical FL algorithms and the published results also show that this scheme preserves model accuracy and reduces the computational cost compared to other secure solutions.

We also continued our efforts from previous tasks based on ensemble learning, where multiple learning algorithms are used to obtain better predictive performance compared to any of the constituent learning algorithms individually. Following the assumption that the goal of any ML problem is to find a single model that best predicts our desired outcome, and since we can often not produce a model that is most accurate in all cases, ensemble methods take a myriad of models into account, and average these models to produce one final model. Thus the common approach to use ensemble learning is to train several models on the same dataset, and aggregate the results using one single ensemble model. In addition to our other implementations we have also followed this approach in collaboration with partners from UC2, the main idea is also related to FL. We have applied an ensemble model to aggregate distributed ML results for predicting/classifying the same problem, trained on different local datasets at servers of the involved parties.

This approach allows parties to collaborate with others in order to jointly solve a problem, without exposing their private data to each other and thus preserving the data privacy. Depending on the parties datasets, and their description, whether they have the same feature set or different feature set, there is a use case where the parties should share their trained model between each other in order to retrain the ensemble model avoiding the need of sharing their data for that purpose. Only the final results of local evaluations are aggregated, the actual training data is not shared with others. We also want to point out that the security guarantees for methods based on data aggregation (ensemble learning, FL), are different compared to encryption methods.

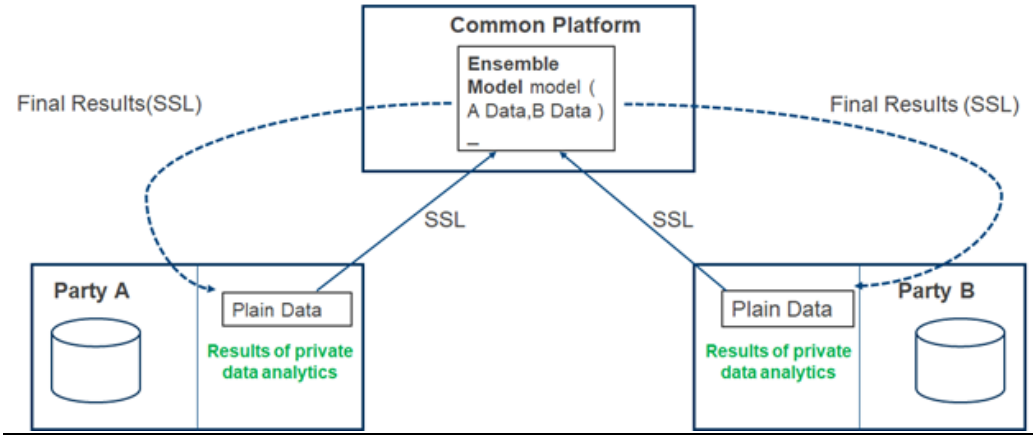

Figure 30: Common Platform for Federated Deep Learning

VFL using SHAP values :

SHAP values interpret the impact of having a certain value for a given feature in comparison to the prediction we'd make if that feature took some baseline value.

The suggested solution provides a capability to run classification ML algorithms over more than one datasets belonging to different and, at times, rival parties. Training is performed without sharing any of the raw data between the various parties, and the final model provides one single prediction while keeping data privacy and security.

The way to withhold these constraints is by running federated ML models, over each of the data sets separately, and then share only the SHAP values generated by each of the models.

The SHAP values from all of the federated ML are used as input to a new classification ML algorithm, which provides a single prediction based only on it incorporating the information from each dataset in the shape of SHAP values.

T4.5 Transformation of algorithms to privacy-preserving certified

Our progress in this task is interlinked with the other tasks in WP4, where on one hand we continued with the development of our privacy-preserving solutions (e.g. our prototype for encrypted TL and the new library for private set intersection, both are outcomes of T4.1 and T4.2) and on the other we also started to work on more secure solutions for FL (T4.4) base for example on multikey HE as already described above.

We also included our PSI library to the software catalog of EUHubs4Data and we wrote a newsletter contribution to inform about the functionalities of the new solution (https://www.trusts-data.eu/private-set-intersection/). We additionally protected the communication channel between the two PSI parties using a TLS connection. For this, we used the *rustls* library, an implementation of TLS in the Rust programming language. Our implementation allows for both self-signed certificates, as well as traditional public-key infrastructure. We also carried out additional intensive benchmarks for one setting where both parties are in the same network (e.g. a datacenter in Frankfurt) and for two further settings where the two parties are in two separate geographical areas (Frankfurt-Paris and Frankfurt-Ohio). The benchmarks were executed on Amazon Web Services (AWS) EC2 servers with both parties running an Ubuntu Server 20.04 LTS image on a c5.xlarge configuration. The following graphic depicts the performance results with respect to the set size of the new PSI library for these settings.

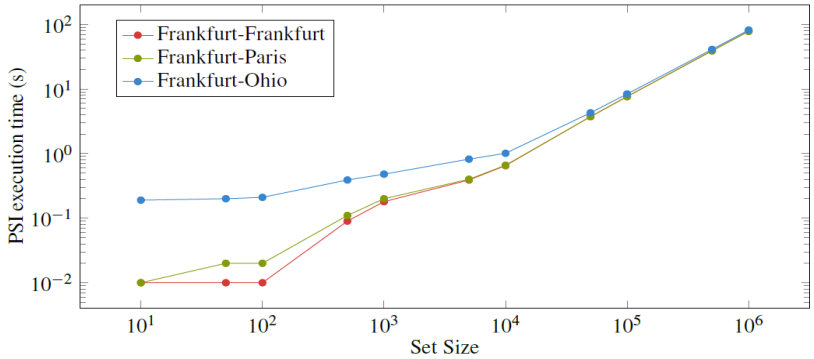

Figure 31: Relation of Performance Results to size of PSI Library

We can see that for small set sizes, the additional latency in the Frankfurt-Ohio scenario leads to an increased runtime, however, this small additional latency is insignificant for larger set sizes and the three different networking scenarios are practically identical in terms of runtime. We have compared these results to the prototype PSI solution that has been developed in the H2020 project SafeDEED, where only benchmarks of small data sets were possible due to the memory usage constraints of the old solution. The new PSI library is usually about 10x times faster than the old one and now also allows the usage of very large data sets. If both parties hold a set of 5 000 items, the new PSI demonstrator takes about 0.39 seconds, while the old SafeDEED solution took 4.54s in the Frankfurt-Frankfurt scenario. The underlying PSI protocol used in the PSI demonstrator is suited to the scenario of imbalanced set sizes, where larger server set sizes are more beneficial for the protocol.

The latest version of our new PSI library is already in use by project partners and has been integrated into the prototypes for UC2 in TRUSTS.

3.4.3 Next Steps

1. WP4 is exploring new VFL methods that will provide results of collaborative analytics while preserving data privacy.
2. finalizing the research regarding SHAP values as input for VFL.
3. will continue with the implementation of our prototypes and carry out a series of tests on benchmark datasets to assure model performance and privacy preservation.
4. will enhance the developed anonymization algorithm and enhance the risk analysis tools by adding anonymization tools.

3.5 WP5 Demonstration of the TRUSTS Platform in 3 business-oriented Use Cases

The WP5, is led by EBOS, and is aiming to **demonstrate the TRUSTS Platform in three business-oriented use cases** which showcase the sharing, trading, (re)use of data and services, and also reporting to the overall TRUSTS objective:

- to demonstrate and realise the potential of the TRUSTS Platform in three UCs targeting the industry sectors of corporate business data in the financial and operator industries while ensuring it is supported by a viable, compliant and impactful governance, legal and business model.

3.5.1 Objectives

According to the GA, the WP5 is focused on:

5. Setting up the test environment and performing the relevant planning and pilot operational management for trials in three pilots;
6. Conducting advanced field trials within the following sectors: Financial Institutions, Telecom Operators, Corporate data providers, etc.;
7. Using the test results and data to deliver impact analysis and impact assessment reports to systematically address the pilots' stakeholder perspectives.

The Table 6 below states the TRUSTS objective 4 that is related to WP5 and the work committed.

Table 6: TRUSTS Objective 4, related to WP5

TRUSTS Objective 4: WP5 Demonstration of the TRUSTS Platform in 3 business oriented use cases
Targeted Effort
To demonstrate the added value of the TRUSTS Platform in 3 business-oriented use cases which showcase the sharing, trading, (re)use of data and services and result in added value generated through innovative applications built on multiple open and proprietary data sources.

Achieving this objective requires the implementation of the three projects use cases:

- **Use Case 1 "The Anti-Money Laundering compliance use case"**: Smart big-data sharing and analytics for Anti-Money Laundering (AML)
- **Use Case 2 "The agile marketing through data correlation use case":** Agile marketing activities through correlation of anonymized banking and operators' data
- **Use Case 3"The data acquisition to improve customer support services use case":** Data processing and visualisation services for Big Financial Data, specifically to advance new ways of human-computer interaction (e.g. chatbots).

The work of WP5 is organised in three tasks as illustrated in the figure 30 below:

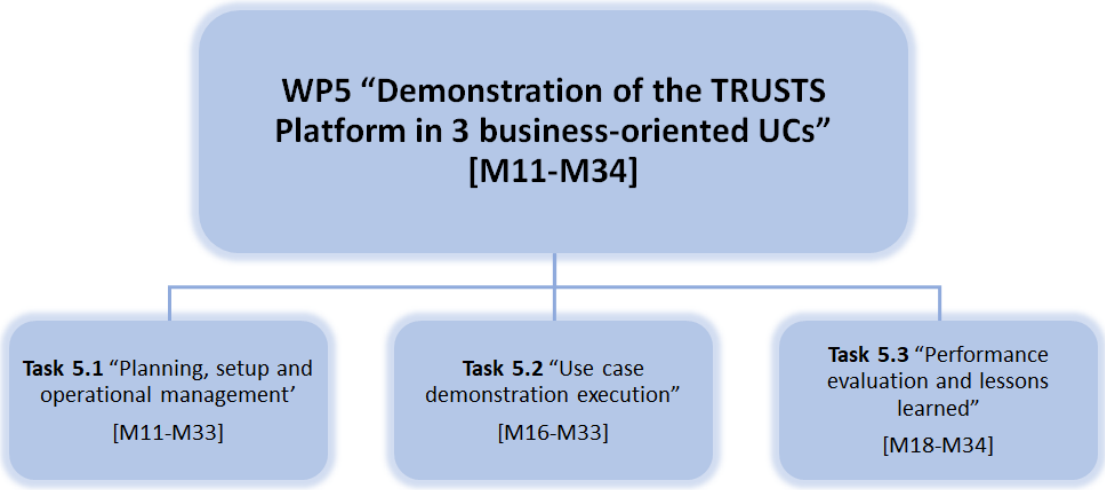

Figure 32: WP5 tasks

During the second year of the TRUSTS project, WP5 primarily focused on setting up the test environment and the relevant planning and operational management for the execution of the first phase of the three UC trials that was set to start by May 2021 until November 2021. The first report under the WP5 was also produced and submitted reporting on the overall plan of the first cycle of the TRUSTS trials.

Subsequently conducting advanced field trials within the sectors of Financial Institutions, Telecom Operators, Corporate data providers, etc., while demonstrating and validating the TRUSTS Platform started in April 2021 as per the WP5 Gantt Chart (see figure 31) created illustrating the WP5 milestones along with the projects commitments and lifecycle. The well-defined UCs effectively adopted in WP5, aim to ensure the

© TRUSTS, 2021

technology innovations in WP3 and WP4 and thoroughly test the TRUSTS solutions and business aspects, involving actors that represent all targeted sectors.

Each of the three use cases has provided measurable KPIs to quantify the overall performance and achieved developments. Based on the three UCs, a requirements elicitation process was set up in collaboration with WP2, to make sure that the TRUSTS results are applicable in day-to-day business.

Regarding GDPR and research ethics compliance, all three UCs have in common that the legal basis for each processing activity was ensured in advance, and data encryption techniques were employed during transit and storage. In addition, the data will only be used for the TRUSTS trials. Milestone 3, the First Pilot Deployment M16 (April 2021), was achieved with – the Final Methodologies for the business validation of use case results; and the Final Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition – in collaboration with WP2 and the first Pilot planning and operational management reports of WP5 documented in D5.1 submitted in March 2021.

3.5.2 Progress achieved

In light of project developments as well as close monitoring of WP5 and T5.1, they both started two months earlier, in M11, November 2020 and not M13, January 2021 as per the GA, due to extensive time calculated for preparatory activities, services implementation and deployment to the TRUSTS Platform prior the start of the trials. With the trials starting M16 (April 2021) and always within project scope, more intensive effort was needed to ensure work of high standards. Along with the services implementation and deployment to the TRUSTS Platform (MVP v.0) and Task 5.2, the UC trials were well equipped. Besides enhanced effort to support the deliverable author, as the first deliverable of WP5, D5.1 was due in M14, required while focusing on projects objectives and commitments, more realistic timelines to accommodate the more intensive workload were created.

During the second year of the project, WP5 focused on the planning of the three UCs for the execution of the first cycle of the trials, the actual field trials execution and the initial lessons learned derived from the first cycle.

In detail:

Task 5.1 “Planning, setup and operational management’

Task 5.1 is led by EBOS, and it is the management and monitoring task of WP5, following and coordinating the WP activities and the planning and preparation of the three UCs trials.

Task 5.1 objective as per the GA, is to provide the necessary demonstration testbench to the stakeholders, so as to be able to demonstrate through actual field trials that the TRUSTS Platform is capable of supporting the KPI requirements defined in WP1. It is to provide the process in planning, setting-up and managing the demonstration pilots and their UCs, so that a constant interaction cycle of progress delivers the results incrementally emerged from the pilots that followed the Deming Plan-Do-Check-Act (PDCA) cycle created under this task so that there will be a constant interplay between their progress and the technical developments.

Specifically, this task prepared a Gantt chart (see Figure 31) of the Tasks and WPs duration as well as the tests to follow throughout the lifetime of the WP, and milestones.

© TRUSTS, 2021

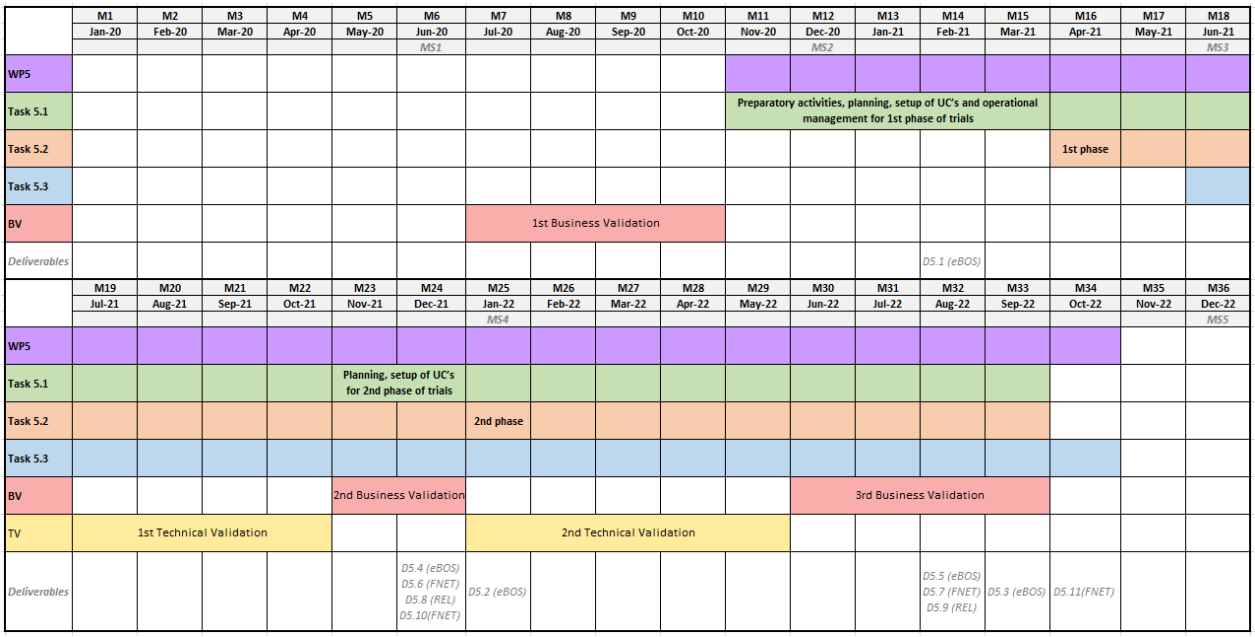

Figure 33: WP5 Gantt Chart

Following the Gantt Chart, Task 5.1 provided the overall planning as well as the setup activities for the deployment and testing of the three UCs for the first cycle of trials, reported in detail in D5.1 (submitted March 2021).

T5.1 also monitored the implementation progress of the execution of the three UCs (started May 2021, under T5.2), in coordination with the involved stakeholders/partners as well as the WP3 and WP4 leaders to guarantee the compliance with the project objectives.

This task also ensured that all activities in the pilots are carried out in accordance with the ethics principles defined in WP6.

The first deliverable of Task 5.1 was submitted in March (M15), titled “D5.1 – Pilot planning and operational management report I'' giving a detailed description of the planning and operational information of the three business-oriented UCs according to the business requirements and FR, as defined in WP2 “Requirements Elicitation and Specification” for the first phase of the trials. The set-up, the procedure for the implementation and testing plan for the UCs which is continuously updated and reported at the end of each demonstration phase, was initially reported in D5.1 for the first phase of the trials. Briefly this deliverable reported the status of WP5 and T5.1 “Planning, setup and operational management”, the preparation of the first phase of the trials and summarized the efforts taken thus far in the interrelated WP2 and WP3. The objective of D5.1 was to define and to document the framework setup for the implementation of the UCs, during the planning phase. These procedures were structured, and validated by all project partners before the actual implementation, in order to check their feasibility and applicability of the UCs and maximize its efficiency. An overall plan of the trials in which these activities took place to achieve their objectives, was done and was followed, by using the Plan-Do-Check-Act cycle model as detailed throughout the report.

Additionally, under Task 5.1, FORTH implemented the Banking Analytics and Insights application, which constitutes a major application for the realization of UC2. The application is responsible for correlating datasets of financial and CRM nature in order to create targeted analyses. Specifically, FORTH designed and developed the frontend/GUI of the application as well as the rest of the application’s architecture. The

© TRUSTS, 2021

application's architecture is built in a loosely coupled manner, where every component is being a standalone entity in a separate docker image.

Following are the components incorporated to the application's architecture:

- The application frontend/GUI
- The Coordinator Server, which is responsible for the necessary backend operations with regard to the coordination of the backend components according to specific workflows (e.g., data ingestion to the Data Management System from files, updating the upload statuses, etc.).
- The Data Management System, comprising the Search Engine, the Ingest Pipeline and GUI for monitoring the Search Engine. It is responsible for the communication between the components (metadata), storing of the datasets destined for analysis as well as the logging of the App's components outputs.

The analytics and insights offered by the application are developed and tailored as per the specifications and needs of PB as well as the data offered by NOVA. In detail, the application combines the datasets provided by NOVA and PB, so as to create insights' and analytics' smartboards pertaining to a diverse range of categories including, telecom contracts and digital data traffic trends per location, loan analytics as well as financial and digital info per geographical area to indicate areas with high commercial interest.

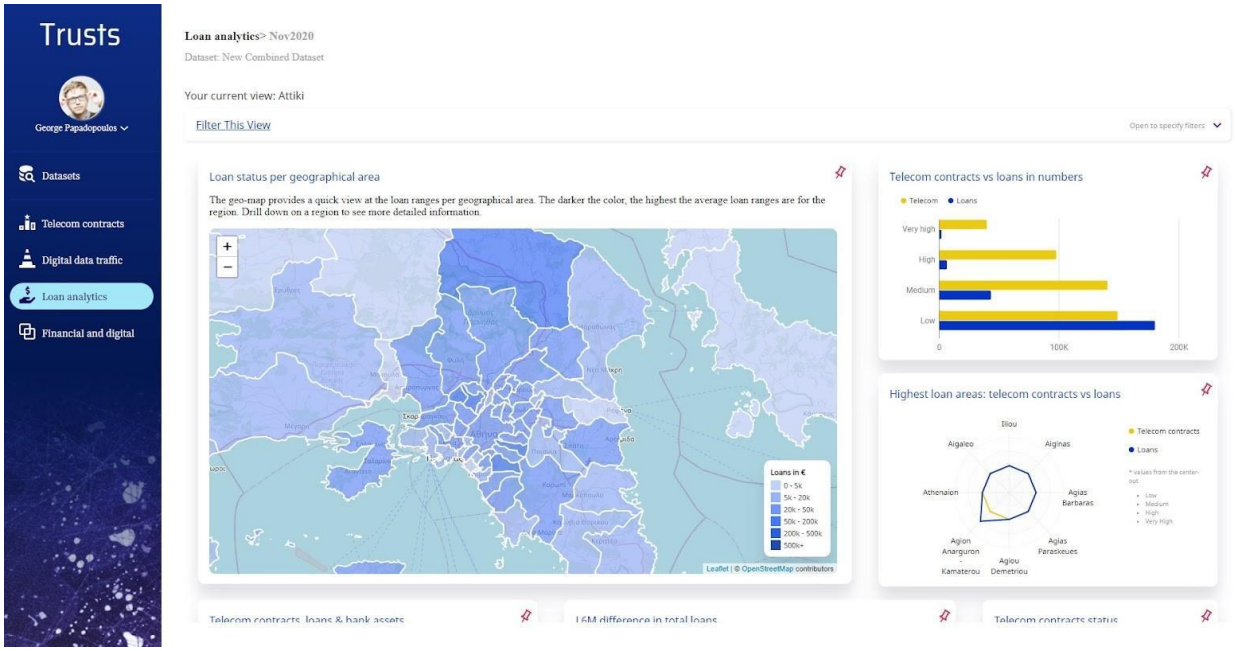

Figure 34: Analytics and Insights – Loan analytics page

TRUSTS envisages to exceed the current data market state of the art, accommodating a full range of data trading and respective collaboration services in one unified platform dealing with both sensitive private and industrial data. These, along with the ability to develop a sustainable business model, are key differentiators of TRUSTS from other data marketplaces.

The process described in D5.1, continued and all the activities, implementation, deployment and testing of the UC's were monitored and will be updated in the second report documenting the planning for the second cycle of the TRUSTS trials commencing January 2022, which is due M25 (in January 2022).

© TRUSTS, 2021

A third and final report, D5.3 will be submitted in M33 (in September 2022) documenting the final outcome of Task 5.1 and the concluded UCs monitoring throughout the lifetime of the project.

Task 5.2 Use case demonstration execution

In this task, led by NOVA (ex. FNET), pilot actors perform actual testing and validation activities in cooperation. The aim is to perform real life trials involving business and technical stakeholders. Deployments of all three UCs aim to imitate the envisaged commercial service thus enabling technical, performance and business validation. This task is divided to three subtasks, each one addressing the trial requirements of the respective UCs, i.e.:

- **ST5.2.1** Smart big-data sharing and analytics for Anti-Money Laundering (AML) compliance
- **ST5.2.2** Agile marketing activities through correlation of anonymized banking and operators' data
- **ST5.2.3** Buying data from a data marketplace to improve Natural Interaction.

Despite the fact that each UC has varying requirements in terms of deployment, processes, datasets and applications usage, T5.2 defined a consistent methodology to archive trials deployment in a consistent, comparable and quantifiable manner. Respective achievements are summarised in the following:

- Based on the T2.2 work which described and matured the use cases scenarios, aiming also at acquiring consensus between all partners and on their final description in the D5.1 deliverable which was produced by T5.1. T5.1 yielded an extensive set of Test Cases per UC scenario. The Test Case template aims at incorporating in a systematic manner all essential aspects of the trials performance and validation. The Test Case template structure was proposed by T5.2 and adopted by T2.3 to constitute an integral part of the testing framework. The test case template is illustrated in the table below:

Table 7: Test Case Template

TRUSTS UC[X]	**UC title**				
SCENARIO UC[X]-SC[Y]	**Scenario title**				
Test Case UC[X]-SC[Y]-TC[Z]	**Test Case title**				
Test Owner	Name, Partner				
Test schedule	Test dates				
Trial Preconditions	Required TRUSTS Platform, required data, required applications, required parameterization, etc.				
Test participants	List of test participants (if necessary they will sign and informed consent)				
Test steps	**KPIs to be tested**	**EXPECTED RESULTS**	**ACTUAL RESULTS**	**PASS/FAIL**	**ADDITIONAL NOTES**

© TRUSTS, 2021

Trial postconditions					
Results evaluation, lessons learned and recommendation	**Business**		**Technical**		
Other supporting material	E.g. Trial recording				

Further to the Test Case description template an end to end process was defined in order to track trials steps, stakeholders participation and evaluation. The key process steps were:

- Detailed infrastructure deployment description per UC
- Definition of a trials registry to serve as a single point of reference to track schedule and results of all trials
- Appointment of UC and trial leaders
- UC trials stakeholders definition. A respective Inform and Consent form has been defined to be signed prior to each trial.
- Trial performance including specific sets of Test Cases
- Stakeholders Questionnaire to be filled right after each trial
- Evidence of trials performance e.g. videos, photos, etc.

The trials designed and executed in the first Cycle of the UC trials were well defined in order to address all aspects of the implementation and produce sound results. In terms of stakeholders that participate in the trials:

- UC1's notion of having ML models sent back to the application provider involves multiple parties, as follows:
    1. EBOS who has 3 different roles, acting as:
       a) Application provider (app1, app2, app3)
       b) Data provider (utilizing RDC database)
       c) End-user by providing input data necessary for the execution of the trials
    2. NOVA acting as an end-user by providing input data necessary for the execution of the trials
    3. InBestMe acting as an end-user by providing input data necessary for the execution of the trials

    All above parties are also involved on the metadata provided to the platform (i.e. trained models).

- UC2 involves five different parties:
    - The data provider/consumer (PB and NOVA),
    - The application providers (KNOW, FORTH, LST),
    - In addition, FORTH and LST play the integrator role for PB and NOVA respectively.

© TRUSTS, 2021

In UC3, an external service provider was involved, with REL acting as an intermediary.

T5.2 aims at providing a 360° evaluation of the TRUSTS Platform. To this end, all platform MVPs (provided by WP3) are evaluated in correlation to the respective Mock-Ups (provided by WP3) aiming at conveying all end product usability aspects even at early implementation stages:

- MVPs are used to provide a bottom up evaluation of the technological performance and processes completion,
- Mock-Ups are used in a top down evaluation demonstrating the end state usability aspect of the aforementioned MVP technological evaluation.

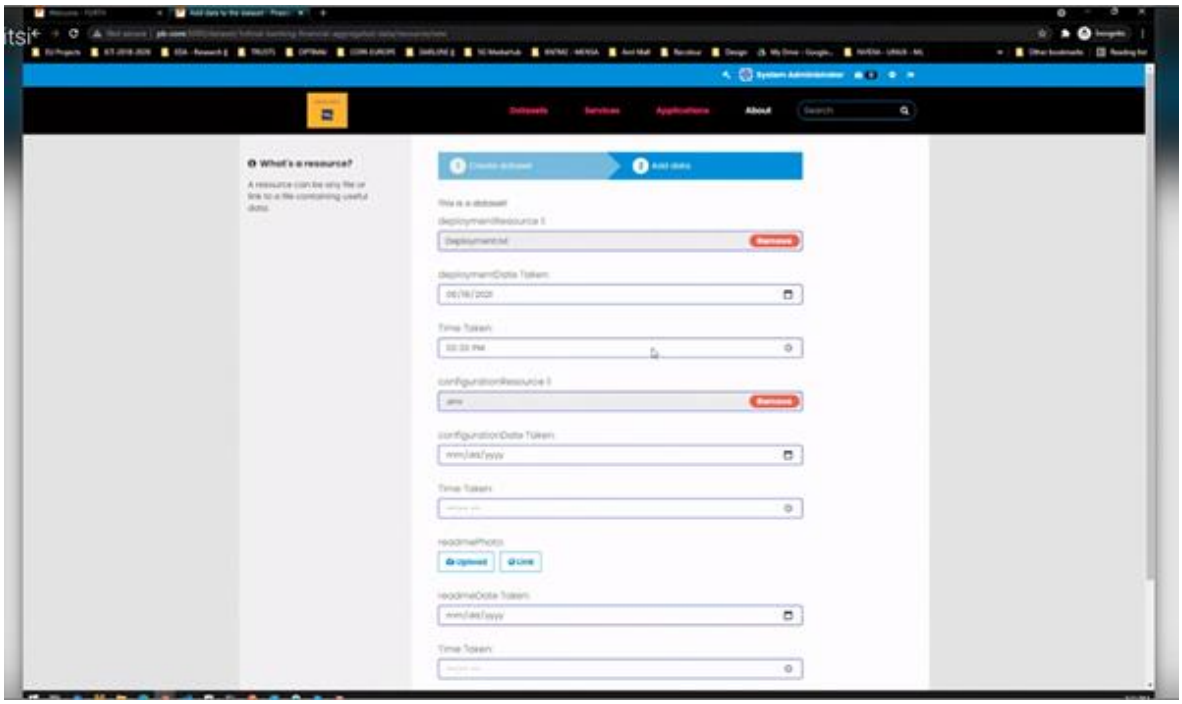

Figure 35: Mock-Ups of the TRUSTS Platform

T5.2 aims to closely collaborate with all other WPs providing timely feedback based on ground truth and findings acquired through the business and technology stakeholders interaction with the platform. Thus T5.2 will evaluate the platform via real life trials rather than laboratory tests. T5.2 performs trials in two cycles: the first ending in M24 and the second in M32 producing two deliverables and providing input to T5.3 which analyses the lessons learned acquired through the UC trials.

© TRUSTS, 2021

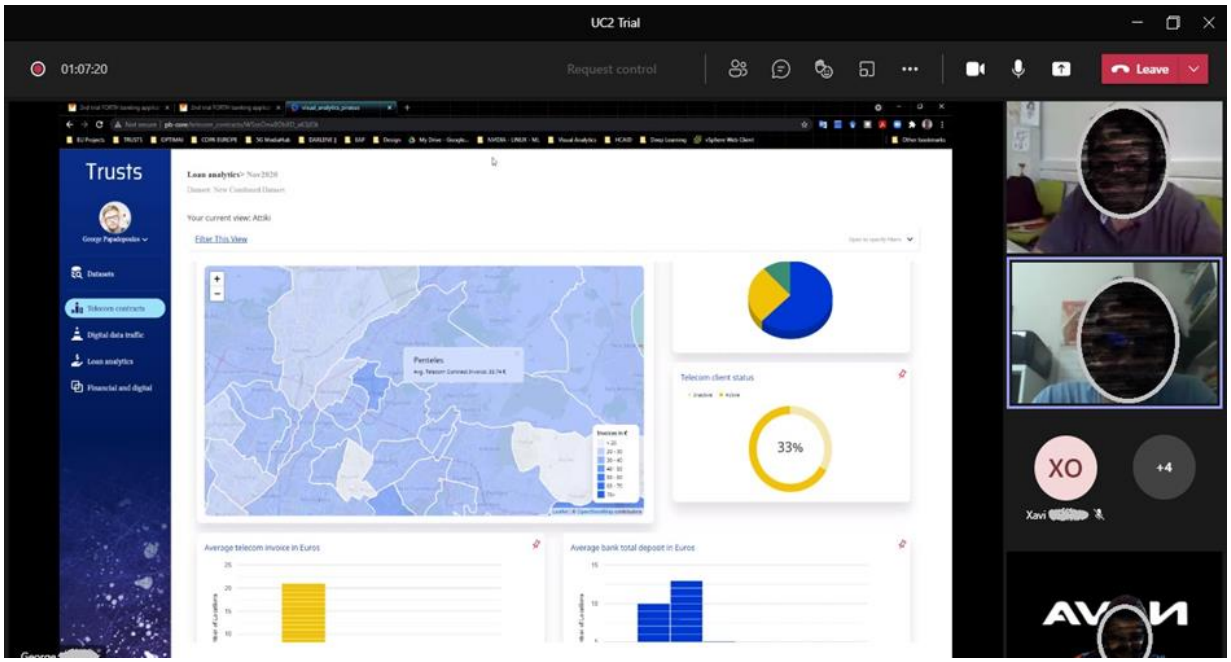

Figure 36: TRUSTS trials in progress (screenshot material #1)

During Cycle 1, 21 trial sessions were performed by the three UCs with 121 Test Cases (*Note: Detailed analysis of the trials is presented in the deliverable D5.10)*. All functionalities implemented in MVP v.1 were thoroughly tested.

In addition, the UI Mock-up was demonstrated in several trials. Significant effort was made to include participants external to the TRUSTS team. All participants provided comments during the trials which were kept by the trial leader for further analysis in the form of questionnaire feedback.

All the information above has continuously been analysed to provide feedback to the platform development work packages (e.g. WP3) and to be used for the Cycle 1 lessons learned report production (D5.10) as well as to provide input for the updated FR list (D2.3) and the Cycle 2 trials planning.

© TRUSTS, 2021

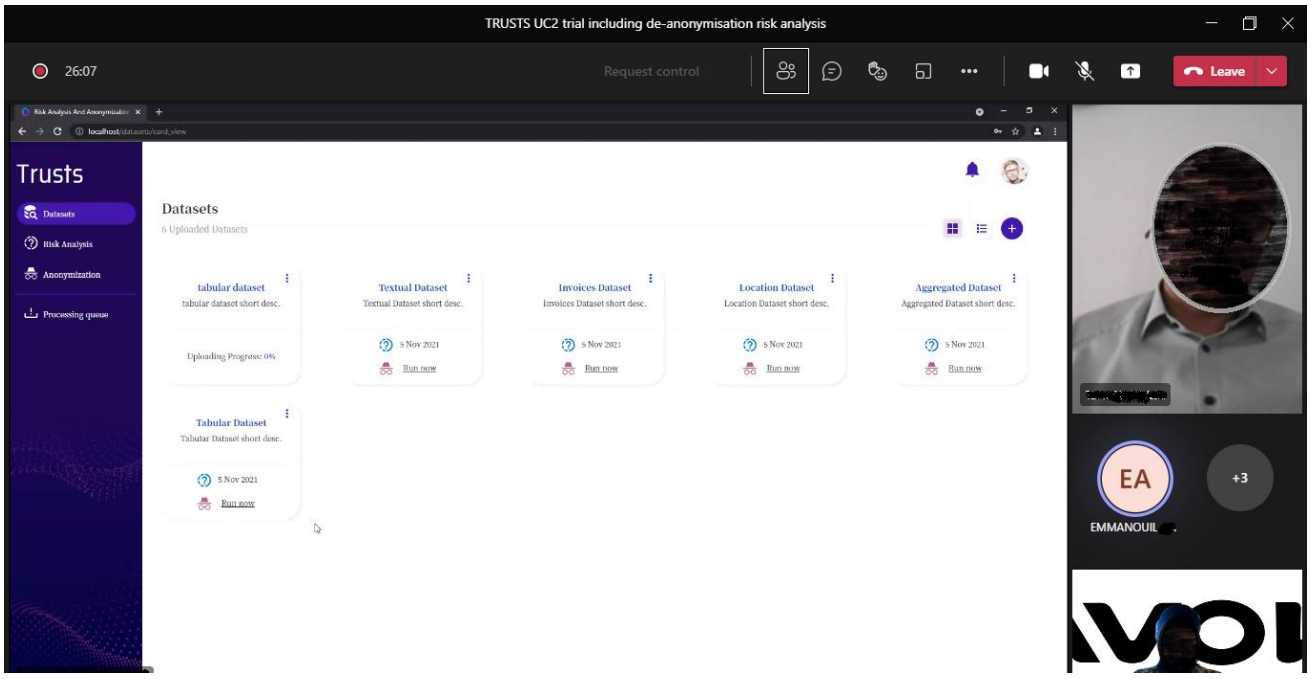

Figure 37: TRUSTS trials in progress (screenshot material #2)

Task 5.3 Performance evaluation and lessons learned

This task commenced in June 2021 (M18). Despite this fact, planning in order to systematically collect and analyze the lessons learned from the trials towards providing consistent and timely feedback has been finalized.

The purpose of this task can be described from the following two aspects:

1. The performance of each use case will be evaluated, particularly from the KPI perspective to illustrate how the TRUSTS Platform capabilities can be leveraged for different applications in each use case and
2. According to the results received from each use case in every agile-based iteration, the task will provide requirements and suggestions to further improve both functional and non-functional capabilities of TRUSTS.

In collaboration with all UC partners, a consistent and unified trials evaluation methodology was defined in order to achieve comparability and complementarity of the analysis results. In brief, the following process is adopted in order to systematically collect information from the trials.

Step 1: MVP1 deployment announcement

- An official announcement for the MVP1 availability is provided by WP3
- Each UC appoints a leader

Step 2: Detailed UC infrastructure description

- Each UC describes the detailed deployment

Step 3: UC trial stakeholders definition

- Each UC defines the stakeholders that participate in the trials.
- All stakeholders prior to each trial sign a respective “Inform & Consent form”.
- Signed forms are kept by the UC leader

Step 4: Trials implementation

- The UC leader appoints a leader for each trial.
- The trial leader defines the test cases that are executed in each trial.
- The trial leader safeguards that all necessary stakeholders participate in the trial and they sign the inform and consent form
- Each stakeholder responds to a questionnaire following each trial.
- The trial leader fills the Test Case forms as defined by the Task 5.2 as well as the trials registry keeping also all the respective information e.g. trial questionnaire responses.
- Evidence should be provided for the performance of the trial e.g. video recording, photos, screenshots, etc.

The trials registry has the following format:

UC leader	

Trials Information					
No	Trial leader	Date	Test Cases	Link to documentation	Comments/Evidence
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Figure 38: UCs trials registry

Lessons learned were derived following the analysis of the Test Case forms, the stakeholders’ questionnaire and any related observations. Beyond the frequent collaboration between the WP5 partners, the TRUSTS consortium was updated for the results at least once a month while respective workshops were planned. In the event of urgent matters, ad hoc collaboration was achieved with the respective WPs.

Initial T5.3 results were reported in D5.10 entitled “Performance evaluation and lessons learned report I” which was in parallel submitted in December 2021 (M24).

The key Lessons Learnt in all three TRUSTS UCs are:

- During Trials, all MVPv.1 functionalities were tested through well-defined scenarios. Usability during the trials was limited due to the fact that technicality was mandatory.
- MVPv.1 User interaction mechanism is not adequate for a contemporary operation marketplace.

© TRUSTS, 2021

- The stakeholders noted that TRUSTS usage documentation is mandatory.
- The search process on a prototype service/dataset catalogue demonstrated convenient filters during these Trials, but there is certainly home for improvement.
- Harvesting of all TRUSTS information in each node may impede scalability.

Key TRUSTS business strategy focus is:

- Become a fully operational European Data Marketplace, providing Intellectual Property management for personal and non-personal related data.
- Act as a platform Federator, laying the groundwork for an ecosystem that will enable federation of independent data marketplaces.
- Create framework conditions to facilitate the emergence of an ecosystem of an ever-increasing number of companies around TRUSTS.

Along these lines, during the Circle 1 Trials, the stakeholders raised respective issues and put forward proposed recommendations. In particular:

- Privacy concerns and fears of disclosure of trade secrets were raised when testing the AML and PSI Apps. The stakeholders suggested:
- TRUSTS current UIs and workflows are not friendly to use, do not follow business logic and are quite restricted.
- Trials should be more complex, resembling the real life usage of TRUSTS.
- Contracting, payment and remuneration processes were not available in the current state of the TRUSTS development.

The above business recommendations are consistent with the FR for the TRUSTS platform that are defined in deliverable D2.2.

3.5.3 Next Steps

The second cycle of the TRUSTS UC trials are set to commence in January 2022. The field trials will proceed as planned with the second trial phase and the lessons learned reported at the end of this phase respectively.

The updated version of the planning and operational management deliverable, D5.2 is planned to be documented and submitted by January 2022, M25.

Nevertheless, an agreement on augmenting the trials in a way that demonstrates the multi-subscribers platform usage, as discussed during the review, could be significant towards proving and disseminating the advanced TRUSTS technological and business proposition. To this end, WP5 aims at adapting the trail design of the subsequent cycles in this direction.

Each of the three business-oriented UCs demonstrates different abilities of the unified operational platform. In WP5 we aim at augmenting the trial scenarios in a way that they clearly demonstrate the platforms' ability to trade data assets, collaborate and fulfil stringent business requirements in a unified environment including a significant number of assets and stakeholders. In this way the ability of the operational platform to support efficiently and in a trustworthy manner multiparty usage towards creating a federated data ecosystem will be clearly evaluated and demonstrated.

© TRUSTS, 2021

Table 8: WP5 Year 2 Deliverables

Deliverable	Task	Title	Author	Due Date	Submitted
D5.1	**T5.1**	Pilot planning and operational management report I	**EBOS**	Feb-2021	**19-Mar-2021**
D5.4	**T5.2**	Actual field trials of Use Case 1 v.1	**EBOS**	Dec-2021	**pending Dec-2021**
D5.6	**T5.2**	Actual field trials of Use Case 2 v.1	**NOVA**	Dec-2021	**pending Dec-2021**
D5.8	**T5.2**	Actual field trials of Use Case 3 v.1	**REL**	Dec-2021	**pending Dec-2021**
D5.10	**T5.3**	Performance evaluation and lessons learned report I	**NOVA**	Dec-2021	**pending Dec-2021**
D5.2	**T5.1**	Pilot planning and operational management report II	**EBOS**	Jan-2022	**pending January 2022**

WP5 concludes in M34, October 2022, executing two cycles of trials producing the lessons learned of the project's achievements and improvements while overall producing a total of eleven (11) deliverables.

3.6 WP6 Legal & Ethical Framework

3.6.1 Objectives

Overall objective of this WP is to provide an analysis of the relevant legal acts and develop a robust legal and ethical framework for the TRUSTS Platform to ensure sustainability and compliance of the innovation brought by the project with all relevant regulations and ethics principles. The main objectives of WP6 are to:

- Provide a set of requirements in order for the project to be carried out in compliance with the principles of research ethics
- Analyse the European laws and regulations relevant to data transactions and the TRUSTS Platform development
- Define a set of legal and ethical requirements and identify potential legal and ethical obstacles
- Generate recommendations for policy makers and stakeholders in the field based on best practices and potential identified gaps

© TRUSTS, 2021

3.6.2 Progress achieved

In WP6 KUL provided an overview of legal frameworks in order for the project to be carried out in compliance with the principles of research ethics; analysed the European laws and regulations relevant to data transactions and the TRUSTS Platform development; defined a set of legal and ethical requirements and identified potential legal and ethical obstacles, generated recommendations for policy makers and stakeholders in the field based on best practices. In the framework of our research we have already submitted two deliverables. Our D6.1 on research ethics provided all information regarding the compliance of the project research activities with research ethics and, in particular, with the H2020 Programme Ethics Guidance. D6.2 on Legal and Ethical requirements studied patchwork of legal frameworks applying to data transactions. KUL also performed research on thinking of data as a commodity that could be turned into a tradable asset. In addition, the analysis was performed on data market ecosystems as those based on the concept called the 'commodification' of data.

In task 6.1 KUL provided all information regarding the compliance of the project research activities with research ethics and, in particular, with the H2020 Programme Ethics Guidance. D6.1 was submitted on 28 February 2021 (M14). In the framework of this task, KUL analysed the research ethics principles in order to address the legal and ethical issues arising from the research activities that will be conducted in the course of the TRUSTS project. The development, testing and validation must comply with ethical principles to respect the individuals involved and to prevent harm.

The deliverable 6.2 identified the relevant EU legal frameworks applicable to various data transactions that are envisaged in TRUSTS. More specifically, it provided insight into the privacy and data protection legal framework supporting the data sharing in compliance with the EU rules. It informed partners on the main concepts of the ePrivacy legal frameworks and their relationship with the GDPR. Furthermore, this deliverable was a continuation of the work done in WP9 with regard to anonymization of personal data. The present deliverable provided further conceptual legal information on privacy preserving techniques and some of techniques that might be relevant for TRUSTS partners.

KUL provided an overview of the role of platforms and/or intermediaries in the field of data sharing (e.g. in the light of the Platform to Business Regulation ('P2B Regulation')). It then analysed EU legislation applicable to data sharing in B2B context such as the regulation of unfair commercial practices between businesses, mainly at national level, taking Germany, France and Belgium as an example. Based on the European Commission 'Guidance on sharing private sector data in the European data economy' of 2018, it provided consortium partners a non-exhaustive list of considerations which may help in the preparation and/or negotiation of data usage agreements. In this task we also discussed the role of data for competition law analysis and the relationship between competition law (Article 101 and Article 102 TFEU) and personal data protection law (such as the GDPR), when personal data is at stake.

It also provided an overview of the relevant regulatory frameworks relating to transactions over financial data. First of all, it covered frameworks aiming at facilitating the fight against anti-money laundering and terrorist financing. Then it looked at Payment Services Directives, their scope of application and their relationship with the GDPR with the objective of informing the consortium partners of potential challenges in an anticipatory manner. This deliverable also presented potential points of contact between the blockchain technology and the operation of the law. The blockchain technology and smart contracts are then subsequently addressed in T3.2. Finally, this deliverable outlined the possible ethical implications of data sharing within the TRUSTS Platform. It offered a high-level analysis of ethical issues in data sharing with the

© TRUSTS, 2021

use of AI-driven tools and provided an overview of the ethics requirements for Trustworthy AI as defined by the High-Level Expert Group (HLEG) in non-binding 'Ethics Guidelines for Trustworthy AI'.

KUL was also involved in several dissemination activities such as Legal aspects of data sharing – a TRUSTS & Safe-DEED Webinar 31 March 2021; Podcast "Data sharing and EU's digital strategic autonomy".

3.6.3 Next Steps

At the moment WP6 is conducting research on new legal frameworks discussed in the EU such as the Data Governance Act, Digital Services Act and Digital Markets Act.

Task 6.3 leads to the deliverable that assesses the legal and ethical considerations generated through the course of the project and especially through the integration of the requirements identified in D6.2 Legal and Ethical Requirements. This task is led by KUL, which is a partner entrusted with the regulatory monitoring of all the legal frameworks relevant for the TRUSTS project. At the moment KUL is focussed on analyzing the Data Governance Act and its impact on the TRUSTS project. In the framework of our analysis, we already issued the White Paper on Data Governance Act[13], which offers an academic perspective to the discussion on the proposal for a Data Governance Act put forward by the European Commission in November 2020.

In task 6.4 KUL will develop recommendations for relevant stakeholders and policy makers, based on the identified legal gaps and lessons learned throughout the running of the project. The work for this task is particularly relevant taking into account the regulatory changes that will take place in the EU data related legal framework, including DSA, DMA and Data Act. The fast evolving EU legal framework calls for the analysis of the intersection between the GDPR and Data Governance Act with respect to definitional, institutional and substantive aspects of those legal regimes.

3.7 WP7 Business Model, Exploitation & Innovation Impact Assurance

3.7.1 Objectives

The objectives of the WP7 "Business Model, Exploitation & Impact Assurance" are to develop a feasible business model to sustain the results of the project, mobilize an ecosystem, and conduct concrete actions for commercializing the data market platform. Thus the WP sets out to conduct market research on what business models for data markets exist around the world. The main focus is on business models combining scientific and non-scientific founders since TRUSTS has the same mixed private and public ownership structure.

The main deliverables during the project will be on the ecosystem and its needs regarding the innovation aspects and intellectual property and data management. The WP will establish pre-conditions for successful business models and best practices. To achieve the same, the WP is divided into the following tasks:

1. T7.1 Sustainable business models
2. T7.2 Developing and structuring the platform engagement
3. T7.3 Intellectual Property and Data Stewardship
4. T7.4 Standardisation uptake and recommendations
5. T7.5 Commercialization initiatives and action plan

[13] Cf. available a thttps://www.researchgate.net/publication/352690055_White_Paper_on_the_Data_Governance_Act.

6. T7.6 Innovation Impact Assurance

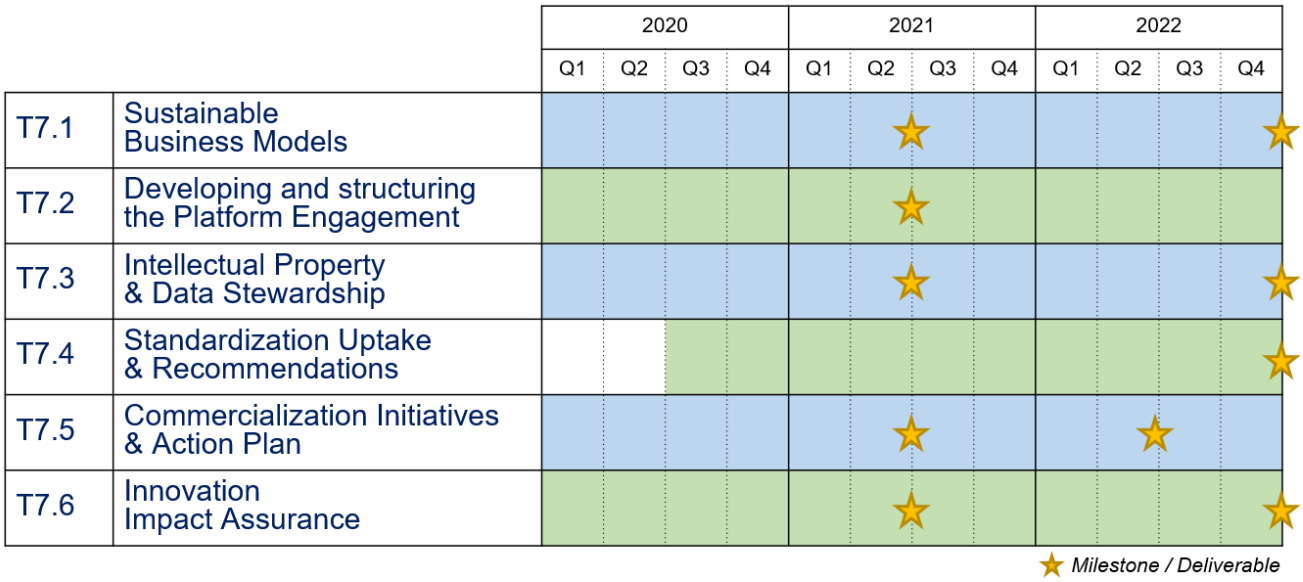

Figure 39: Milestones and Deliverables of WP7

All work tasks have been fully activated and are delivering against the formal deliverable structure as per the DoA and the above table.

3.7.2 Progress Achieved

Overview

The work in WP7 (Sustainable Business Models, Exploitation, and Innovation Impact Assurance) has been reported in D7.1, D7.3, D7.4, D7.7 and D7.9. Research in data markets and data market federation (T7.1) yielded a business-model centric, unified taxonomy which forms the basis of identifying and selecting viable business model options. This is aligned and builds on the wider market study conducted in T2.1. The approach to exploitation and commercialization (T7.5) has been drafted, and is pending detailing based on aforementioned business model options and research into product-market-fit, utilizing the stakeholder engagement strategy and plan (T7.2) which has been finalized and is being implemented. To inform auxiliary services of a future data market operator, and to ease data preparation particularly for onboarding of data sellers, mechanisms for DS and IPR protection were drafted (T7.6). Lastly, WP7 has complemented the administrative project management approach of WP1, in a series of targeted interventions to help alignment and focus across all project work packages, as part of innovation impact assurance (T7.6).

T7.1 Sustainable business models

The aim of this task is to select a viable, feasible and sustainable business model for the data marketplace platform developed in the project. To inform the business model development, first, through desk research and interviews, a range of potential data marketplace business models were explored, leading to taxonomies

© TRUSTS, 2021

of possible business model design options. The taxonomy development was the primary focus of the first half of the project.

The primary function of the developed taxonomies is to:

1. Contextualize and position TRUSTS within the developed taxonomies, and
2. Explore the potential of business models for TRUSTS.

A significant highlight of these taxonomies is to emphasize TRUSTS' roles in the EU data economy, which goes beyond that of a 'basic' data marketplace: TRUSTS will also be a federator and an ecosystem facilitator of data marketplaces. Thus, the business model taxonomies were developed considering these roles.

In total, **four business model taxonomies** were developed. The first two taxonomies specifically explore business models of data marketplaces and build on desk research. Whereas the first taxonomy considers data marketplaces that are not specific for a given industry, the second taxonomy explores the automotive industry. The third taxonomy is created concerning the TRUSTS role as a federator and an ecosystem facilitator of data marketplaces. Because a taxonomy is generally developed based on characteristics of existing phenomena, the development of the third taxonomy was complemented with expert opinion from workshop participants with expertise in data marketplaces, business models, and technical requirements. Finally, a unified taxonomy was developed to contextualize the previous three taxonomies for TRUSTS' needs (See Figure 37). The example of the unified taxonomy is presented in Figure 38.

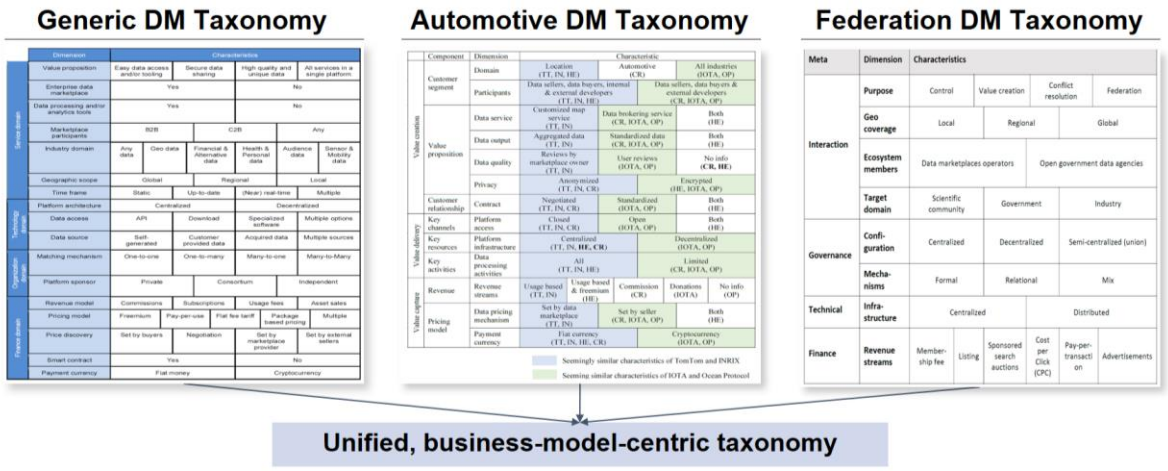

Figure 40: Two data market taxonomies and a federator taxonomy as inputs to create a unified taxonomy

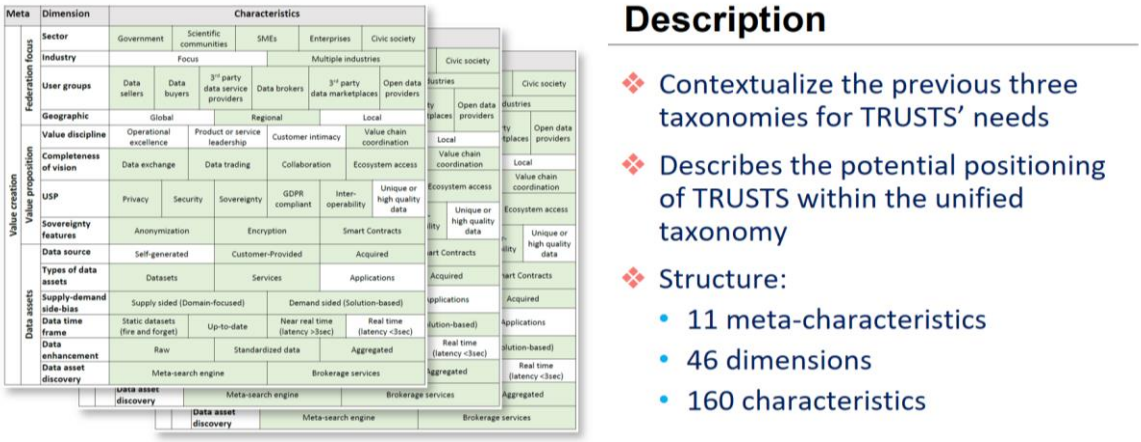

Figure 41: A Unified, business-model-centric taxonomy

The third and unified taxonomy development is informed by workshops organized with technical experts and business actors. Accordingly, the taxonomies developed in this report are empirically and theoretically informed grounded in an understanding of practical considerations of the envisaged roles for a sustainable data marketplace.

After describing the taxonomies, T7.1 also presents risks, opportunities, and business requirements that TRUSTS should consider. For example, TRUSTS can consider value propositions that offer a solution focus instead of raw data trading, the needs of strong customer relationships, and the seeding strategy for attracting new end users. Considering TRUSTS role as a federator of data marketplaces, the main challenges related to perceived insufficient value creation and perceived risk & cost should be considered in future business model development efforts. Nevertheless, this role opens seven new business model opportunities to be considered, such as providing a one-stop-shop via a standardized portal, providing commissioned brokerage for data buyers who look for solution-based data assets, and establishing shared services for non-differentiating capabilities (e.g., billing) and others. As an ecosystem facilitator of data marketplaces, TRUSTS needs to reflect on the mentioned challenges, such as lack of shared visions across the ecosystem members, low level of goal congruence among actors in the ecosystem, growing complexities, and others. Nonetheless, these roles open business model opportunities for TRUSTS.

T7.1 "Sustainable business models" will continue to work towards the second half of the project phase to develop a business model for TRUSTS. The focus will be on selecting business models based on the insight extracted from this deliverable. The business model will be developed by applying tools for business model innovation as developed in TUD's award winning platform businessmakeover.eu. The tools will be applied in workshops with project participants and, later on with outside stakeholders to validate hypotheses and to stress test the business models options. After developing the business models, the evaluation will be done in three ways:

1. by conducting a summative evaluation on the implications of business model choices on critical success factors that measure the viability of the business model;
2. by informing T7.5 on concrete actions and activities needed to realize the business model and testing the feasibility of these actions based on T7.5 findings;
3. by applying TUD's method of business model stress-testing to evaluate the sustainability of the business models in different future scenarios (e.g., different levels of citizen trust in data economy or different levels of regulatory regimes).

© TRUSTS, 2021

T7.2 Developing and structuring the platform engagement

This task is elaborating and tailoring the Stakeholder Engagement Approach for the TRUSTS project in order to foster the community around TRUSTS and its results. The main aim is to ensure a commercial uptake of the project outcome by a vivid and informed community. Therefore, an analysis of the TRUSTS stakeholder landscape has been conducted within this task as the basis for the engagement approach, followed by the derivation of a concrete Stakeholder Engagement Plan (SEP) that proposes specific activities to engage stakeholders at relevant times.

The identification and mapping of the relevant stakeholders for TRUSTS has been used as a starting point and lead to several user groups, organisations or individuals identified which has been divided into the following five main stakeholder categories: TRUSTS platform user – Customers; TRUSTS platform users – Technology/Infrastructure Operators and Providers; Associations, Organizations & Initiatives; Research & Academy; EC & Policy-makers. Subsequently, a structured prioritization of the stakeholder categories has been conducted, considering the estimated effort that is needed to engage them as well as the potential impact of their engagement on the project's result (see Table 9 "TRUSTS Stakeholder Categories prioritisation" below).

Table 9: TRUSTS Stakeholder Categories prioritisation

Stakeholder Category	Effort	Impact
TRUSTS Platform Users: Customers	*	***
TRUSTS Platform Users: Technology and infrastructure providers	***	***
Associations/Organisations and initiatives	*	**
Research and Academy	*	**
EC & Policy-makers	**	*
Legend	** Little effort* **** High effort*	**Little impact* **** Big impact*

Since the prioritization revealed the high potential of TRUSTS platform users, as well in the form of future customers as in the form of TRUSTS technology and infrastructure providers, those two types of TRUSTS Platform Customers have been picked up to be detailed as well as analysed with regard to their impact on the projects' results. In general, the analysis showed that commercialization efforts pursued in TRUSTS' task T7.5 need to consider the dynamics between different users and that the data market should stimulate both supply and demand.

The above-mentioned results have been then utilized by transferring them into suitable strategies for the TRUSTS Stakeholder Engagement. This stakeholder engagement approach takes into account three building

© TRUSTS, 2021

blocks to balance the strategy (also depicted in Figure 39: “TRUSTS Stakeholder Engagement Building Blocks”):

1. Stakeholder engagement to achieve project core objectives
2. Stakeholder engagement for access to cross-project advisory and multiplicators
3. Stakeholder engagement to support work package delivery

Figure 42: TRUSTS Stakeholder Engagement Building Blocks

As a last step, the Stakeholder Engagement Strategy has been transferred into tangible actions in the form of a Stakeholder Engagement Plan (SEP) to be carried out by the TRUSTS project tasks and partners and which is complementary to the outreach activities coordinated by the project's communication team (WP8). These actionable recommendations aim at establishing a vibrant and sustainable community around TRUSTS and even beyond its lifetime and need to be implemented and realized as a next step.

Already several activities have been carried out – for instance the Data Market Dialogue "TRUSTS Datamarket: Use Cases in Trustless Collaboration and Data Sharing", where additional requirements from potential users have been collected directly[14].

T7.3 Intellectual Property and Data Stewardship

In task T7.4, we target challenges around Intellectual Property Rights (IPR) and Data Stewardship (DS). The objective is to protect original data owners/providers and resellers of enriched data whilst supporting innovation and value extraction. Obviously, bare minimum legal requirements have to be reflected in the technical design of the TRUSTS platform as well as in the general terms and governing contracts. This must be complemented with effective mechanisms to report and address suspected IPR infringement. But beyond, TRUSTS has to define its overall approach as to how active its role should be in the domain of IPR protection, and – within legal confinements – where to strike the balance between opposing interests of different TRUSTS user groups vis-à-vis a sustainably viable business model. Particularly for SMEs, regulations and (dispositive) rights regarding the use and re-use of their IPR is not self-evident. The same holds true for requirements towards SMEs acting as buyer of data for aggregation, enrichment, and onward sales.

[14] Link to the event (11/03/21): https://hopin.com/events/workshop-trusts-datamarket-use-cases-in-secure-data-collaboration-and-sharing

© TRUSTS, 2021

Task T7.4 defines, how TRUSTS will go about related segmentation of user groups (if any), and different onboarding as well as continuous information/education requirements and services. In turn, this links to enabling DS on the side of (prospective) data providers. Existing attempts of data markets have often suffered from the lack of available data and data quality, because many organizations - in particular SMEs and semi-governmental agencies do not have a sufficient internal data governance, and do not “know what they know” or how to commercialize this data in a meaningful, yet protected way that also has them retain control over their data integrity. This task will research the support services requirements for different (potential) data provider groups to optimize eased attraction and onboarding of (SME) data providers onto the platform, to enable value creation and extraction within TRUSTS.

To date, the first version of the report “Supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship I'' (deliverable D7.4) has been created based on desk research, expert interviews and two workshops with constituents of the former Data Market Austria (DMA) project and the Austrian Data Intelligence Offensive (DIO). The purpose of this deliverable is to set up the guidelines on how IPR will be managed by the TRUSTS consortium and will be continuously updated and reported at the end of the project by M36 as D7.5 “Supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship II”. It outlines:

1. Legal requirements to be embedded in the platform's terms of use,
2. Defined mechanisms to report suspected Intellectual Property (IP) infringement,
3. Proposed onboarding IPR protection information and education requirements for TRUSTS user groups, and
4. Proposed DS Support Services for different (potential) data provider groups to optimize eased attraction and onboarding of Small Medium Enterprises (SMEs) data providers.

The following technical measures have been identified:

1. Securing the IPR both physically and digitally by the use of cryptography
2. Data anonymization by removing personal or confidential data field before publication
3. FL as a technique for decentralized learning where private and sensitive data never have to leave their local storage location.
4. Ensemble learning, where aggregated data sets are used.

Further, the IDS, on which TRUSTS is building on, is offering as mechanisms to support the IPR protection the IDS metadata broker, as an intermediate relevant for TRUSTS IPR mapping, as well as providing access and usage control, and the IDS Clearing House as a monitoring instance for transactions and indicator for fair use.

© TRUSTS, 2021

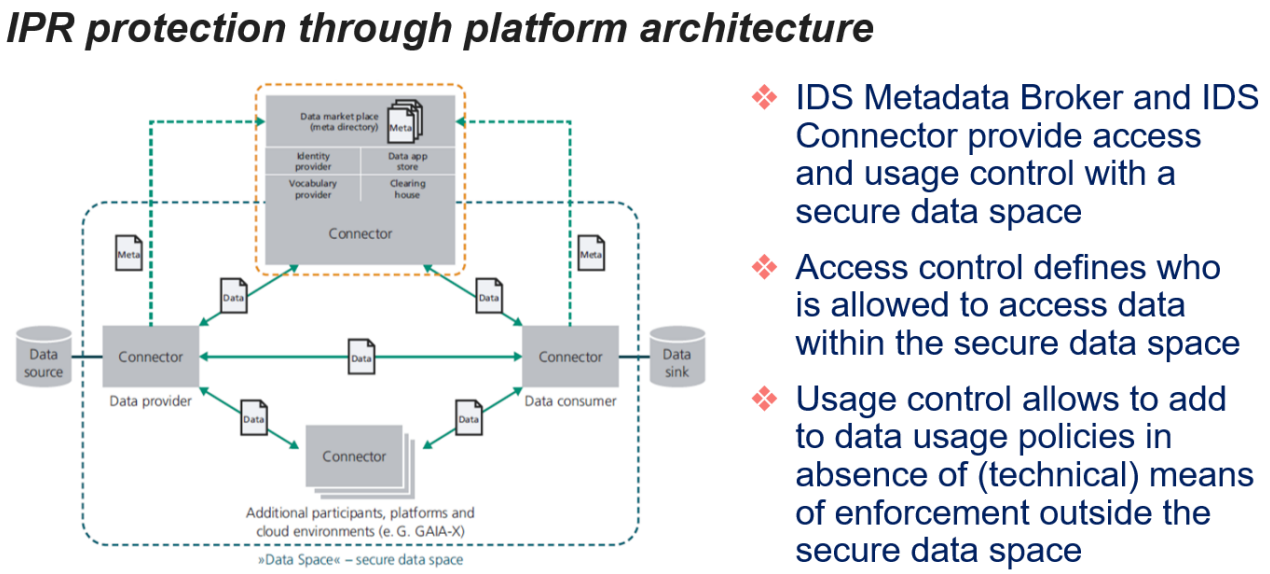

Figure 43: IPR Protection through Platform Architecture

The following contractual measures for IPR protection have been drafted:

DRAFT "CODE OF CONDUCT FOR USING THE TRUSTS PLATFORM" (CC)
PRELIMINARY REMARKS / PREAMBLE
DRAFT §1 GENERAL PRINCIPLES
DRAFT §2 GENERAL RULES OF CONDUCT - RESPECT / DISCRIMINATION
DRAFT §3 CONFLICTS OF INTEREST
DRAFT §4 DATA PROTECTION / CONFIDENTIALITY
DRAFT §6 VIOLATIONS AND SANCTIONS
DRAFT „TERMS AND CONDITIONS FOR USING TRUSTS SERVICES" (TC)
CREATION OF THIS DRAFT TERMS AND CONDITIONS (T&C)
DRAFT §1) DEFINITIONS
DRAFT §2) SCOPE OF THE TERMS & CONDITIONS
DRAFT §3) THE OPERATOR: THE TRUSTS OPERATING COMPANY (TRUSTS OPCO)
DRAFT §4) TRADING SYSTEM AND CURRENCY
DRAFT §5) GENERAL DUTIES TO COOPERATE
DRAFT §6) PARTICIPATION IN DATA TRADING / LISTING PROCESS
DRAFT §7) DATA TRADE, DATA TRANSMISSION AND ARCHIVING
DRAFT §8) FEES FOR THE USE OF THE TRADING PLATFORM TRUSTS
DRAFT §9) SANCTIONS AND TERMINATION
DRAFT §10) DISPUTE RESOLUTION PROCEDURE
DRAFT §11) MISCELLANEOUS

1. "Code of Conduct for using the TRUSTS Platform" (CC), a framework for amicable cooperation, that will be eventually enriched with sanctions and penalties in the next deliverable of this task if needed, and
2. "Terms and Conditions for using TRUSTS Services" (TC), summarizing the results developed in the project and providing a legal framework for further work on and with TRUSTS.

The task already concluded on several recommendations towards the conceptualization and development of the TRUSTS platform aiming at efficient and affordable IPR protection mechanisms:

1. Secure and legally compliant exchange of the datasets and services,
2. Review of published data to make informed decisions on buying legitimate products,
3. Need for mechanisms that ensure the validity of the datasets and services onboarding process,
4. Users' reputation schemes that should also be supported as a protection measure,
5. Deployment of effective and secure user management, and
6. Provision of inherent protection of private datasets.

From an IPR point of view, three fundamental aspects were identified as important for the further development of the TRUSTS platform and should be considered by a future TRUSTS operator:

1. Cross-system mapping of data assets

© TRUSTS, 2021

2. Actualisation of meta-data from decentralised data storages and data networks
3. Interaction of automatic digital contracts and data assets

Task T7.3 will focus its activities for the remainder of the project on:

1. Refinement of the general concept for dealing with IPR in data markets
2. Support of the implementation of technical IPR measures during development of the platform
3. Finalization the contractual measures (for platform users and consortium members)
4. Drafting of framework conditions for a TRUSTS operating company
5. Preparation of an agreement between the TRUSTS consortium partners, balancing recognition of their IPR and commercial interests of a future operator

T7.4 Standardisation uptake and recommendations

The task deals with the use of standards as well as the suggestions for required extensions of standards in the TRUSTS environment, whereby the main focus is on interoperability and relevant standards in this area. The general goal is to lower current barriers of data sharing given by the multitude of differing approaches to share or exchange data in a trusted and secure way. Therefore, this task is currently elaborating, assessing, and identifying how relevant concepts or standards can and may be developed or extended in order to ensure a state-of-the-art TRUSTS ecosystem. This is happening by having a closer look at what different standards or concepts have been used or considered within TRUSTS already and the collection of qualitative feedback about the relevance and eventually undertaken adaptations. The task includes 3 workshops on practical challenges and perspectives regarding IPR protection and DS.

Based on this, the task will formulate some recommendations about the ones considered as “most promising” and will bring them up for discussion to different standardization bodies and relevant stakeholders like the BDVA. The results will be made available as a report at the end of the project, in December 2022. The Report outlines (1) legal requirements, (2) mechanisms to report IP infringement, (3) onboarding IPR protection information and education requirements, (4) proposed DS support services.

T7.5 Commercialization activities and action plan

This task is advancing by analysing the business targets, pricing models, and operational costs in order to transform the TRUSTS platform into a sustainable ecosystem. Our aim is to explore how TRUSTS platform can successfully enter the European market and achieve profitability, while overcoming potential limitations.

The starting points are evolving, and the core innovations are being monitored with the objective to define concrete actions for the implementation plan. For this commercialisation process a key step is to identify the value proposition that differentiates TRUSTS in front of competitors and strategy. The main added value elements identified in TRUSTS proposal and which constitute its value proposition are:

1. TRUSTS is to be a cross organisational and cross sectorial environment where big data value technologies, and novel data applications can be tested, piloted and exploited,
2. The platform will provide services around the complete data cycle, and components to create data spaces for the creation of intelligent applications.
3. It is being aligned with GAIA-X, IDSA, which are currently reference positions.
4. It will be interoperable, the corporate node will allow organisations to integrate their own infrastructure, to exploit own and shared data, and

© TRUSTS, 2021

5. TRUSTS users in the marketplace will be able to monetise datasets as Data products and Application services.

Even if T7.5 focuses on the TRUSTS platform as a KER, other KERs based on the TRUSTS building blocks (such as services, products, methodologies) will be further identified.

The ecosystem considered is composed of a series of actors with specific functions as:

- Data providers (bringing data),
- Data brokers (communicating data offer and data demand),
- Data scientist (deploying digital technologies),
- Business domain experts (trading-off and linking services),
- Data consumers, and
- Data operators (to maintain the infrastructure, to manage the data usage, accelerators of the data driven innovation).

Once we segment the potential stakeholders, we will propose business models for the different target groups and their needs (data vendors, data buyers and data brokers). To define areas of potential higher revenue, and to structure the results of the Project to be sustainable.

T7.6 Innovation Impact Assurance

Aspiration of this EU funded programme is to bridge from research to market, that is to move beyond (research) output to outcomes and defined impacts. Thus, research findings, concepts and prototypes shall be usable for the next level of development and adoption by pertinent (industry) players in the wider business ecosystem. To enable this, task T7.6 shall work with all work packages and tasks to both ascertain from the outset and throughout the project that the aspired Outcomes and Impacts are kept in mind and guide Output creation. The previous experience from the Data Market Austria shows that this dramatically improves research transferability and hence business viability, whilst it also reduces efforts for conceptualization or technology development. Thus, continuous interactions with all WPs and tasks through regular check-ins, coordinated with project management (WP1) are of paramount importance. In doing so, T7.6 also complements and enriches WP1 by enabling a firmer content-involved challenger role of project management as compared to a more coordinating role.

Agreed deliverable of T7.6 is a continuous interaction with all WPs and tasks, acting as a cross-function to the program to ascertain and optimize innovation impact. An “Innovation Impact Assurance”. Progress made is as follows:

- Ascertained linkage and increased synergies between tasks T2.1, T7.1, T7.5
- Acted as link pin to task T2.4, aligning emerging business architecture and technical architecture. Regular participation in technical conference calls and working sessions
- Conducted a range of evangelist interventions to focus and align all consortia partners to intra-ecosystem interoperability and TRUSTS architecture as both, data market and data market federator

For M18, T7.6 created its interim “Innovation Impact Assurance I” activity report summarizing continuous interactions with all work packages and tasks, acting as a cross-function to the program to ascertain and optimize innovation impact. The report describes the synergistic interlock and delivery of traditional,

© TRUSTS, 2021

administrative project management with agile interventions, and outlines the assessed core challenges to achieving innovation impact in M19-M36.

3.7.3 Next Steps

Work package WP7 will continue to deliver along and towards its defined deliverables outlined in section "Objectives", with a particular focus on enabling a consultative process in developing and reviewing the set of upcoming final deliverables due by June 2022 and December 2022:

- Task T7.1: Final version of the "Sustainable business model for TRUSTS data marketplace" report describing the designed business model to sustain TRUSTS after the project ends. It will build on the taxonomy for data marketplace business models to create, validate, and decide on viable business model options.
- Task T7.2: Support and coordination of delivery against the M18 "Communities Engagement Strategy" report, describing our strategy to widen the community around the platform and how to attract new stakeholders during the project and beyond its lifetime.
- Task T7.3: Final version of the "Supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship" report, utilizing desk research and insights from a workshop with stakeholders of the data economy on their practical challenges with DS and IPR protection. It should be noted that the Data Market Austria is not operational, and thus we continue to substitute for its originally assumed stakeholders.
- Task T7.4: The "Standardization Activities" report will present our recommendations about standardization, especially in the field of our use cases. It will also include our exchanges with standardization bodies during our 2 workshops in 2022.
- Task T7.5: Final version of the "Business plan and Implementation plan" report, describing the strategy of the consortium to transform the platform into a sustainable ecosystem and the TRUSTS business plan (business target, services, pricing, costs, remuneration of partners, etc.).
- Task T7.6: Final version of the "Innovation Impact Assurance" report will summarize continuous interactions with all work packages and tasks, acting as a cross-function to the program to ascertain and optimize innovation impact

The WP7 is currently conducting the detailed planning and preparations to enter 2022 with the following foci:

- Research, drafting, consultation and revision, and finalization of all elements required for the work task reports due by June 2022 and December 2022
- Continuation of external communication and ramp-up of engaging outreach activities
- Delivery of an expert panel, jointly with task T3.3, on an outlook of interoperability between data spaces, federation of data markets, and future scenarios pertaining to the data ecosystem
- Delivery of a final workshop with stakeholders on challenges with data stewardship and IPR protection
- Delivery of workshops with standardization bodies
- Continuous monitoring the validation of the 3 use cases from the business perspective
- Preparation of an agreement between the TRUSTS consortium partners, balancing recognition of their IPR and commercial interests of a future operator
- Definition of potential remuneration models for partners
- Drafting of framework conditions for a TRUSTS operating company

© TRUSTS, 2021

3.8 WP8 Dissemination, Communication & Community Building

3.8.1 Objectives

This WP comprises dissemination, communication, and community building. Ensuring efficient internal and external communication, if it does not fall under the tasks of WP1, is the overall objective of WP8. Efficient communication within the project is achieved through regular virtual meetings where reporting of WP8 measures is shared and further planning is discussed. Communication outside the project is achieved through the quarterly newsletter and the design, production and publication of regular online content (via Website, Social Media, Social Microlearning, YouTube, Podcast-Tool) to inform stakeholders and a wider public at national, European and international level about the project objectives and results. The existing channels also serve the purpose of community building.

Another objective is to ensure open access to (non-confidential) research results and to make sure that these results can be securely accessed and preserved beyond the duration of the project. The project website has been expanded to include a research section to make it easier to read and find the research papers. To make the most out of the results, bring them closer to the stakeholders and promote (science) related skills a training and capacity building programme is created. A concept for this was developed in 2021 (D8.6).

At the end of each project year, WP8 collects and documents the dissemination activities of all partners and reports on them in the annual dissemination report (D8.3 and D8.4).

3.8.2 Progress achieved

This part is an excerpt of the Annual Dissemination Report I (D8.4), which was elaborated and published by WP8 in a detailed manner. The Annual Dissemination Report can be found on the TRUSTS website.

Within 2021 the basis for project communications from the previous year was expanded and used effectively. The media mix was optimized and diversified. Content generation was strengthened through more project output and diversely placed in the media landscape.

T8.1 Dissemination and Communication Strategy, design guide, materials, and communication channels

The communication and dissemination work performed in WP8 follows the general outline of the GA and the D8.1 Dissemination and Communication Strategy, Design Guide, Materials, Communication Channels.

Via the various formats used within WP8, the visibility of the project was strengthened throughout the second project year. Those formats include blog posts, interviews, podcasts, webinars; additionally, project partners attended key events to foster TRUSTS' visibility in an interpersonal manner.

T8.2 'Visual identity, website, and promotional materials'

A coherent and consistent recognition of the project is indispensable for a holistic success of the H2020 project. Within every communication action the in 2020 defined branding and visual identity of the TRUSTS project has been respected.

The basic website was set up prior to the beginning of the project in September 2019. It represents the main communication channel of the project. Within 2021 it was regularly updated and filled with new content – e.g. blogposts, whitepapers, podcasts and research papers. Besides, the website sections were extended. A podcast section was added, as well as a section for webinars and training (in preparation for T8.4) and an own section for research papers.

© TRUSTS, 2021

The upcoming steps include an extra section for the Stakeholder Advisory Board.

T8.3 'Large scale dissemination of projects impacts and results'

A key component of performing communication and dissemination activities was continuity. It is crucial to have contents published on a regular basis. This is why the newsletter and the podcast (with small deviations) were published on a quarterly basis. Blogposts had a higher frequency and Social Media postings were used three times a week to keep attracting stakeholders' attention. According to the ethics guidelines of an EU project, TRUSTS has a clear focus on cooperation. Therefore, in addition to Safe-DEED, other projects were approached by TRUSTS for collaboration (i3-market, DOME 4.0, Kraken). Furthermore, there was an exchange with key initiatives like the BDVA/DAIRO and GAIA-X AISBL.

In terms of KPIs for communication and dissemination, TRUSTS performed very positively and overachieved many of the KPIs by the end of the second year.

Table 10: Impact of communication and dissemination activities

Channel	KPI and estimated number of persons reached	Type of audience reached in the context of the dissemination & communication activities
Project website (including blog posts on news page)	44.900 visits/month 5.300 visitors/month	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology platform, data market standardisation body.
Social media (Twitter, LinkedIn, YouTube ResearchGate)	Twitter: 369 follower, 118 tweets LinkedIn: 379 follower, 128 posts YouTube: 28 subscribers, 10 posts ResearchGate: 8 follower, 12 updates	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology platform, data market standardisation body.
Scientific publications	6 publications	Scientific community (researchers, universities, etc.), policy makers, EU projects, media representatives
Conference attendances	4 attendances	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology platform, data market standardisation body.
Meet-up attendances	14 attendances	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology

© TRUSTS, 2021

		platform, data market standardisation body.
Press releases	1 press release: 9,000 editors; 21,000 mail subscribers[15] 13 visits on website	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology platform, data market standardisation body.
Newsletters	3 newsletters, 2 special issues 891 subscribers **Opening Rate** 2nd NL (addendum 2020): 15,41% 3rd NL: 13,51% 4th NL: 14,56% 5th NL: 17,48% SI 1: 14,68% SI 2: 17,20%	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology platform, data market standardisation body.
Podcasts	3 podcasts 2nd podcast, 14.12.2020: 152 views (YouTube, Website, Podigee) 3rd podcast: 291 views (YouTube, Website, Podigee) 4th podcast: 195 views (YouTube, Website, Podigee) 5th podcast: 37 views (YouTube, Website, Podigee)	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology platform, data market standardisation body.
Webinars	4 webinars, 90 participants/viewers	Scientific community, media representatives, policy makers, data and service providers, data consumers, EU projects, competence center/digital innovation hub, technology platform, data market standardisation body.

T8.4 'Training and capacity building programme'

This task started in M13 at the beginning of the year, led by Relational SA (REL) with support from DIO, SWC, G1, and NOVA (ex. FNET). It involves a training programme and e-learning materials for various stakeholders that could potentially adopt the TRUSTS data market platform within their organizations. The result of "D8.6

[15] https://apa.at/produkt/ots-verbreiten/

– Concept for training and capacity building program," is a summary of training and capacity building plans, if known, their effectiveness.

Existing materials will also be included in the capacity building plan. The aforementioned capacity-building program tries to broaden the playing field of Data Market owners by introducing powerful arguments for the target audiences. In addition to the overview of promising capacity-building programs, the task is also looking at the functioning of an organization as a whole entity.

3.8.3 Next Steps

In the following last year of the project implementation, the communication efforts will be intensified once again in order to solidly anchor the substantial results in the European data community as well as in the general public in the last phase of the project.

In terms of stakeholders engagement activities will strengthen the collaborative work with our SAB and try to engage as proactively as possible with the European Data community. A delayed ”Mid-Term” Event in Q1 2022 will help to reach out to our community with tangible results and progress, followed by a Finale Event in Q4 2022, which will consolidate that engagement. If the situation concerning the pandemic allows it, more live attendances at events to present TRUSTS are planned for the TRUSTS consortium. The Workshop formats will be optimized from a webinar character to a hands-on character (e.g. Tech Tool presentation for anonymization). DIO has already generated an internal roadmap for the last year which will be the guideline for the work carried out.

Additionally, the training and capacity building programme will be implemented by REL and the whole WP8.

The range of the TRUSTS channels will also be additionally strengthened through stronger stakeholder engagement (together with WP1 and WP7), the involvement of the SAB and specific campaigns.

The public outcomes and activities of the project will continue to be published on the project’s website and on Open Access databases (scientific articles) on a regular basis depending on the progress of the project. The project website itself will be developed further, e.g. through integrating a better overview of the deliverables or more blog posts that are TRUSTS related but not necessarily dealing only with TRUSTS itself. In doing that it is possible to create even more awareness of the issue through the dissemination of information.

3.9 WP9 Ethics requirements

3.9.1 Objectives

WP9 “Ethics Requirements” is a WP added by the EC to ensure compliance with the ethics requirements described in the GA, Annex 1. Ethics deliverables shall be considered as a consistent set of measures aimed at ensuring compliance with ethics requirements within the project. Finally, compliance with ethics and legal requirements are considered as a continued effort by the partners, to be continuously maintained.

The concrete objectives consist of a set of measures that were undertaken in order to ensure the compliance with data protection principles. Specifically, we did the following:

- The templates of the informed consent/assent forms and information sheets (in language and terms intelligible to the participants) were kept on file.

© TRUSTS, 2021

- Prepared a statement from the designated Data Protection Officer that all personal data collection and processing will be carried out according to EU and national legislation.
- Explained how all of the data they intend to process are relevant and limited to the purposes of the research project (in accordance with the 'data minimisation 'principle).
- Described the anonymisation/pseudonymisation techniques that will be implemented must be submitted as a deliverable.
- Provided an explanation how the data subjects will be informed of the existence of the profiling/ tracking, its possible consequences and how their fundamental rights will be safeguarded. This must be submitted as a deliverable. The beneficiary must provide details on the Artificial Intelligence/ Data Mining system and related decision making procedures including information about human actors' roles and responsibilities.
- Described a set of precautions to eliminate or mitigate potential algorithmic biases and explain how the model will be able to justify the results it has provided for specific situations. This must be submitted as a deliverable.
- Confirmed that the beneficiary has a lawful basis for the data processing and that the appropriate technical and organisational measures are in place to safeguard the rights of the data subjects must be submitted as a deliverable.

3.9.2 Progress achieved

The goal of WP9 is to ensure compliance with the ethics requirements described in the GA, in particular in the DoA part B. In order to ensure compliance of TRUSTS project with "ethics requirements" described in the GA, the following process has been set up:

- A questionnaire was drawn and circulated amongst all TRUSTS partners (27 March 2020). It is accompanied by a background note providing further explanation on applicable data protection legal provisions in order to ease the filling of the questionnaire. The filling of the questionnaire shall be used case-specific.
- 3 questionnaires were drawn, namely one questionnaire for every use case as coordinated by respective use case leaders.
- A virtual meeting was convened (15 April 2020) to have a general discussion between partners on the ethics requirements and on how to comply with them.
- A virtual meeting was convened for every use case (21 April and 24 April 2020) in order to tailor the ethics deliverables.
- Based on the information gathered through the questionnaires and the virtual meetings, a first version of the ethics deliverables was drawn and circulated amongst the partners (18 May 2020). After internal review amongst TRUSTS partners, the ethics deliverables were submitted to the EC.

An Ethics Screening took place on 16 July 2019 resulting in six pre-grant requirements related to the recruitment of research participants and the protection of personal data and ten post-grant ethics requirements. The pre-grant requirements for POPD from the Ethics Screening were already satisfied during the GA signature (14 October 2019). The EthSR noted that the project is 'headed in the right direction with regards to ethics', but an ethics check was necessary to review the required submissions and DMP. The EC identified some remaining post-grant requirements that are set for review at the Ethics Check. The Ethics Review (Report) took place in December 2020. To address the concerns raised in the Ethics Review, consent forms were updated as requested by the Commission, further information was collected on ML/AI and anonymisation. All ethics requirements were submitted as deliverables and can be consulted by the Commission at any moment. The findings of the Ethics Check Report from December 2020 are listed below:

© TRUSTS, 2021

Table 11: Overview of EC Ethics Requirements

Humans	
D9.1: H – Requirement No. 3	Resubmitted to the Commission with all the questions addressed and consent forms updates and data processing operations explained
Protection of Personal Data	
D9.4: POPD – Requirement No. 7	Resubmitted to the Commission with all the questions addressed and consent forms updated and data processing operations explained
D9.5: POPD – Requirement No. 10	Resubmitted to the Commission with all the questions addressed and techniques for compliance with data minimization and purpose limitation principles explained by all the partners
D9.8: POPD – Requirement No. 14	Resubmitted to the Commission with all the questions addressed and additional information on the potential use of profiling provided
Other Ethics Issues	
D9.9: OEI – Requirement No. 15	Resubmitted to the Commission with all the questions addressed and additional information on AI and ML provided by partners

Based on these results, D9.1, D9.4, D9.5, D9.8 and D9.9 – which were considered by reviewers as partially fulfilled and open for monitoring – have been updated and re-submitted as one overarching deliverable D9.9 in March 2021. The respective contents from D9.5 and D9.8 have been added as annexes for D9.9.

3.9.3 Next Steps

Ethics deliverables shall be considered as a consistent set of measures aimed at ensuring compliance with ethics requirements within the project. Ethics compliance is a constant follow up work which is being performed by all the partners throughout the project, we consider this as an ongoing work item. Next steps take place around some future work items, e.g. continuous monitoring of legal and ethical compliance, KUL being the central contact point for ethics related questions and ongoing close collaboration between KUL and LUH and partner's legal departments and DPOs.

© TRUSTS, 2021

4 Progress within specific Leads

4.1 Scientific Lead

Objectives
The key objectives of scientific coordination in TRUSTS are to: - Ensure compliance with the H2020 Open Access policy. - Promote the application of good research practices. - Monitor and foster progress towards collaborative research within the project. - Create and enforce the structure for reporting on scientific progress. - Provide opportunities for identifying and/or enhancing research opportunities. - Ensure research activities remain focused on and cover the Call's specific challenges and objectives. - Facilitate learning of research lessons at a metalevel - Maximise research synergies and opportunities
Progress achieved
The key objectives of Scientific Coordination in TRUSTS remain the same as in the previous year's iteration of this report. At the beginning of the reporting period, the consortium partners were already aware of the reporting structures and expectations regarding publication practices, their reporting rules and coordination and these have been followed. Within the reporting period, the consortium was successful in getting 7 research papers accepted, bringing the overall number of research papers resulting from the TRUSTS project to 9. All of these papers provided by the consortium have been made available as Open Access in line with the H2020 Open Access mandate. Ten more research papers are planned by the consortium in Year 3 of the project. This means that we can forecast 19 research papers overall. Out of these ten papers, one has been already accepted with minor revisions, two have been submitted and are under evaluation, two are ready for submission, three are in the making and the work on two more is yet to start. Out of the published papers, four of the seven specific challenge areas (C1: Lack of trusted and secure platforms for sharing personal and industrial data – 5 papers, C2: Lack of privacy-aware analytics methods for secure sharing of personal data and industrial data – 3 papers, C3: Lack of ICT and Data skills seriously limits the capacity of Europe to respond to the digitization challenge – 2 papers, and C5: IT standardisation faces new challenges as technologies converge and federated systems arise, creating new gaps in interoperability – 4 papers) have already been addressed by several publications[16]. While it might seem that more focus is needed on challenges C6 and C7 (C6: Advance the state of the art w.r.t. scalability, computational efficiency of methods to secure desired levels of privacy of personal data and/or confidentiality of commercial data, and C7: Analyse and address privacy/confidentiality threat models and/or incentive models for the sharing of data assets), there are four planned papers addressing

[16] Note that one paper can accept several specific areas of the challenge.

C6 and two planned papers for C7. It is therefore just C4 (Involving SMEs and giving them access to data and technology) which will require more focus of the consortium in the next year.

The Scientific coordinator has been monitoring the research progress and informing the project leadership about its state. The Scientific Coordinator was also encouraging the consortium to work towards research publications in project meetings and his presentations to the consortium.

Next Steps

- Monitoring progress and discussing with relevant consortium partners how to make sure that all of the specific challenges of the Call are addressed in our research. Focus is needed particularly on area C4.
- Further strengthening links between the more research oriented and development oriented work packages (primarily WP3 and WP4).
- Creation of a TRUSTS meta-overview paper, summarising the key research contribution from across the TRUSTS research activities towards the end of Year 3.
- Organise a workshop/forum for TRUSTS researchers as an additional opportunity to identify research synergies in Year 3, fostering particularly cross-partner cooperation and inter-disciplinary collaboration within the project.

4.2 Technical Lead

Objectives

The main objectives of the technical lead are:

2. Architecture design in accordance with task T2.4. **(TL-O1)**
3. Coordination of architectural alignment between work packages 3, 4 and 5. **(TL-O2)**
4. Project management and oversight to execute WP3 according to DoA. **(TL-O3)**

Progress achieved

In the second year of the project, the main goal of the technical lead has been to iterate on the established consensus amongst the technical project partners about the high level architectural aspects of the project. In particular, this consensus is most important for cross-cutting aspects of the project, such as the architectural paradigm, the reuse of existing software and non-FR of strategic importance.

Towards **objective TL-O1 (architecture design)**, the initial version of the TRUSTS platform architecture has been iterated and improved based on feedback from the technical and non-technical partners in the project. The architecture represents the conceptual foundation for the implementation of the TRUSTS platform. In addition, the architecture also allows the project partners with a non-technical view to contribute with cross-cutting requirements of strategic importance.

With regards to objective **TL-O2 (coordination of architectural alignment in WP3/4/5)**, the technical lead has organised or initiated several activities on different scales within the project:

- Project wide activities:

© TRUSTS, 2021

- Participating in the online plenaries on year 2;
- Work package specific activities:
  - Technical sessions in WP3 with external software owners from the IDS ecosystem;
  - Regular meetings on alignment of WP3 with the use case trials in WP5.
  - Continuous tracking of a coarse-task backlog of services that need to be developed de novo and adaptations that have to be done to existing ones.
- Coordination sessions:
  - With leaders of use case trials from WP5.
  - With individual representatives of organisations involved in shaping, designing and implementing the platform.

With regards to objective **TL-O3 (project management and oversight of WP3)**, the technical lead has established the following management instruments:

- A regular telephone conference for all WP3 participants (every two weeks). This allows task leaders to report on progress and obstacles, and provides an easy way for participants to reach out beyond the context of their own task without formal overhead to allow synchronisation between tasks.
- A regular telephone conference for all participants of the architecture design task T2.4, which takes place in alternating weeks to the WP3 telephone conference. It provides the technical participants an avenue to focus on the architectural design, without excluding technical expertise required to anchor the abstract architectural perspective in the technical details.
- A regular telephone conference (every week during development sprints) for coordinating the development sprints of the platform development, which brings together developers and technical experts from the work packages involved in the development sprints.
- A mailing list for WP3 participants to communicate without adding noise to the general project mailing list.
- Availability for telephone calls to exchange feedback and ideas on short notice whenever urgently required by a project participant.

In addition, the technical lead has participated in dissemination and outreach activities (cf. WP8):

- Web presentation: TRUSTS Technical Overview for IDSA Summit, June 2021.

Next Steps

In year 3, towards **objective TL-O1 (architecture design),** if necessary, the architecture specification will be iterated again based on feedback from the UC partners and from the non-technical project partners.

The refinement of the architecture is also intertwined with the planned activities towards **objective TL-O2 (coordination of architectural alignment in WP3/4/5).** In particular, close interaction with the use case trials in WP5 and the integration of the privacy enhancing technologies from WP4 into the platform are expected to require a high degree of effort in year 3.

With regards to objective **TL-O3 (project management and oversight of WP3)**, all related activities from the first year and the second year will be continued and refined. If necessary, additional management

© TRUSTS, 2021

instruments will be instated.

Finally, participation in dissemination and outreach activities continues to be a high priority.

4.3 Innovation Lead

Objectives

Innovation management ensures that the development of both market and technical problems will be accomplished during the project, while enabling the successful implementation of appropriate creative ideas, so that new and improved products, services and processes will belong to the project's output ensuring thus its sustainable update beyond its duration.

Progress achieved

In the second year, TRUSTS started to implement the innovation process agreed within the consortium:

Step	Description
1	Initial TRUSTS innovations declaration
2	Initial market & business model analysis & opportunities (business target, services, pricing, costs, remuneration of partners, etc.)
3	Definition of the initial TRUSTS proposition, competition analysis & USPs. Also define required certifications and key standards compliance
4	Agreement between the partners on the commercialisation alternatives that will be sought
5	Definition of initial potential investors list and the respective investment value proposition
6	Commercialisation actions plan If possible the BP and the commercialisation plan could be validated by external business consultant. Actions can be assisted by such a consultant as well. Commercialisation actions outcome should be reported on a monthly basis
7	Update of TRUSTS innovations, IPR protection actions e.g. patents, etc., standardisation actions, legislation promotion actions, etc.
8	Final business plan, value proposition and commercialisation actions definition

Figure 44: Steps for the Implementation of the Innovation Process

In detail the methodology adopted and actions performed are summarised in the following:

- Safeguard that in each deliverable we identify the innovation topics and if appropriate report then in a respective section according to the innovation focus areas and targets
- All TRUSTS partners introduce their proposals to the TRUSTS innovation registry

The process above led to the adequate response to the Innovation Radar while innovation elements will constitute the "Unique Selling Points" while finalising the WP7 business model.

Next Steps

In the final year of the project, we will focus on:

a. Finalising the business and remuneration models introducing the identified innovation
b. Promote actions towards increasing commercialisation potential according to the process defined above.

© TRUSTS, 2021

4.4 Security Lead

Objectives
The TRUSTS project should ensure data security by carefully evaluated for two main aspects: 1. GDPR Compliance – TRUSTS aims to create a GDPR-compliant European Data Marketplace by respecting the rules and principles through constant guidance from a leading expert partner in law and ethics in the Consortium, KUL, and by taking into account already existing national and international initiatives. 2. Privacy Preserving capabilities – The TRUSTS partners will develop and improve privacy preserving technologies to foster the European Data Economy and at the same time provide business and ethical/legal tools to make these technologies easily adoptable and sustainable. Under these two main pillars, we are keeping track of the project's advancement while ensuring that outcomes are compliant and safe from a security point of view.
Progress achieved
Currently, WP3, WP4 & WP6 have advanced in the security comprehension, where we better understand how the deliverables and outcomes should address security issues and potential vulnerabilities under different scenarios. As a conclusion, the progress of the security-related issues goes according to plan.
Next Steps
While the solution matures and gets ready for implementation, we should keep track of the ethics, GDPR and privacy preserving modules and approach.
Objectives
Overall objective of the legal and ethical lead is to provide an analysis of the relevant legal acts and develop a robust legal and ethical framework for the TRUSTS Platform to ensure sustainability and compliance of the innovation brought by the project with all relevant regulations and ethics principles. The main objectives of WP6 are to: ● Provide a set of requirements in order for the project to be carried out in compliance with the principles of research ethics ● Analyse the European laws and regulations relevant to data transactions and the TRUSTS Platform development ● Define a set of legal and ethical requirements and identify potential legal and ethical obstacles ● Generate recommendations for policy makers and stakeholders in the field based on best practices and potential identified gaps
Progress achieved
Legal lead provided an overview of legal frameworks in order for the project to be carried out in compliance with the principles of research ethics; analysed the European laws and regulations relevant to

data transactions and the TRUSTS Platform development; defined a set of legal and ethical requirements and identified potential legal and ethical obstacles, generated recommendations for policy makers and stakeholders in the field based on best practices. In the framework of our research we have already submitted two deliverables. Our D6.1 on research ethics provided all information regarding the compliance of the project research activities with research ethics and, in particular, with the H2020 Programme Ethics Guidance. D6.2 on Legal and Ethical requirements studied patchwork of legal frameworks applying to data transactions. KUL also performed research on thinking of data as a commodity that could be turned into a tradable asset. In addition, the analysis was performed on data market ecosystems as those based on the concept called the 'commodification' of data.

Next Steps

Next steps will include research on new legal frameworks discussed in the EU such as the Data Governance Act, Digital Services Act and Digital Markets Act.

KUL is focussed on analyzing the Data Governance Act and its impact on the TRUSTS project. In the framework of our analysis, we already issued the White Paper on Data Governance Act, which offers an academic perspective to the discussion on the proposal for a Data Governance Act put forward by the EC in November 2020.

KUL will develop recommendations for relevant stakeholders and policy makers, based on the identified legal gaps and lessons learned throughout the running of the project. The work for this task is particularly relevant taking into account the regulatory changes that will take place in the EU data related legal framework, including DSA, DMA and Data Act. The fast evolving EU legal framework calls for the analysis of the intersection between the GDPR and Data Governance Act with respect to definitional, institutional and substantive aspects of those legal regimes.

4.5 Communication & Community Building Lead

Objectives

The main objectives of WP8 are to ensure efficient internal and external communication and dissemination of the project. In 2021 the basis set in the previous year has been extended and utilized. Channels were optimized and regularly filled with contents. 2021 followed a similar approach as the previous year to guarantee project communications consistency. The project's stakeholders have been identified by T7.2, additional engagement will be outcarried. Online activities and their reception (followers, reactions, etc. on social media platforms) is constantly rising. WP8 supported establishing an external and independent Stakeholder Advisory Board (SAB, lead: WP1), and will continue to engage with the SAB.

Progress achieved

Within 2021 the basis for project communications from the previous year was expanded and used effectively. The media mix was optimized and diversified. Content generation was strengthened through more project output and diversely placed in the media landscape. Those formats include blog posts,

© TRUSTS, 2021

interviews, podcasts, webinars; additionally, project partners attended key events to foster TRUSTS' visibility in an interpersonal manner. This year's communication focus was more on the project's progress, building on last year's general positioning of TRUSTS. The overall plan of WP8 was three-staged: general positioning, progress reports and the promotion of the (realistically expected) outcomes in the last project year. The KPIs defined in D8.1 were mostly overachieved, those not achieved so far will be compensated in 2022.

Next Steps

The conclusions, which build the basis for internal recommendations for TRUSTS' future communication and dissemination activities, can be summarized in same two central pillars as in the first project year:

Ongoing proactivity: Do good and talk about it. The approach of WP8 for communications and dissemination is to share all outputs of the project transparently. The TRUSTS consortium considers all outputs relevant for various sub-groups of the European Data Community. The range of the TRUSTS channels will be strengthened through stronger stakeholder engagement (together with WP1 and WP7), the involvement of the Stakeholder Advisory Board and specific campaigns. Besides, public outcomes and activities of the project will continue to be published on the TRUSTS channels.

Network strengthening: After building first coalitions with other European research and innovation programs (Safe-DEED, DOME4.0, i3-market, etc.), TRUSTS team will continue to do so. Additionally, media connections for content placements like interviews will be aimed at. We will strengthen the collaborative work with our SAB and try to engage as proactively as possible with the European Data community. Events and workshops are planned for 2022, live and/or online.

Additionally, the training and capacity building programme will be implemented by REL and WP8.

4.6 Business & Exploitation Lead

Objectives

Ultimate goal of business & exploitation planning as well as innovation impact assurance is to create a sustainable business model and plan (incl. products & service portfolio, clear SLAs, pricing and billing etc.) for the TRUSTS Platform supported by a wide-reaching Data Innovation Environment. In doing so, it has to be ascertained that the governance and business model for the data marketplace adheres to (privacy) regulation, provides smart contracting to assure quality and service levels, and incentivizes providers, users and owners of data.

Achieving this objective will require to:

A. Analyze the market and build a community of stakeholders (SMEs, start-ups, large enterprises, academics, public administration) around the Data-Services Ecosystem that operate in a clearly regulated environment using innovative business models that ensure the long-term sustainability of the Data-Services Ecosystem beyond the end of the funding period.

© TRUSTS, 2021

B. Undertake standardization activities in view of supporting the standardization of data sharing platform
C. Undertake innovation and IPR management activities
D. Develop an impactful, realistic yet ambitious business plan for paving the way to the commercialisation of the TRUSTS Platform

Progress achieved

Work within the specific work tasks of work package WP7 "Business model, exploitation & innovation impact assurance" is on track (see section on work package WP7, above). However, the role of the Business and Exploitation Lead goes beyond the coordination of the work package activities. The project addresses the specific challenge by focusing on two complementary aspects of the problems hampering the growth of the data economy:

1. Technology: by providing a new state-of-the-art for specific challenges such as that of a secure platform, vitally needed for different data providers to interact confidently and successfully in a market; and
2. Business (commercial, operational, legal, standardization): by developing and testing innovative business models and the effects of current, and future regulations, as threats and incentives for data enterprises.

Therefore, we took an active role in the project to mitigate risks inherent in the dichotomy of R&D project activities on the one side with their focus on value creation in specific domains, and the project aspiration to deliver a commercially sustainable data market embedded in a vibrant ecosystem, with a focus on adequate value capture, on the other side.

<u>Project-wide activities:</u>

- Participation and contribution to Management Board meetings, as per the project governance structure
- Participation and contribution in the project kick-off and all consortia online plenaries, with a particular focus on aligning all consortia partners on the project mandate and pertinent topics arising from work package WP7, e.g. Lessons learned from DMA, Platform federation, TRUSTS business model, and the Data market platform operator

Figure 45: Various presentations on TRUSTS' Innovation Impact Assurance and Business Model Considerations

© TRUSTS, 2021

- Participation and contribution to monthly consortia calls
- Organisation of a multitude of internal workshops, e.g. All Hands for “Positioning of TRUSTS” and “TRUSTS taxonomy”

Figure 46: Presentations of “TRUSTS WP BusTech Alignment” and “Positioning of TRUSTS in the European data economy”

Work package specific activities:

- For detail: See section on WP7

Bilateral coordination and ideation sessions:

- Regular interaction with prior DMA consortia partners, and stakeholders of the Austrian data landscape
- Close coordination with and direct support of task leader 2.1 “EU and worldwide data market” in task planning and tactics ideation, thereby increasing project-relevance of market research and strengthening the linkage to task 7.1
- Close coordination with and direct support of task leader 2.2 “Industry-specific functional requirements elicitation and analysis” in the solicitation of user requirements towards the platform
- Close coordination with the Technical Lead and regular participation in the weekly calls of task T2.4 “Architecture design and technical specifications”, thereby tapping into the planning of the technical implementation in work package WP3.
- Close coordination with the task leader T3.3 “Interoperability Solutions” to jointly advocate interoperability beyond demonstration of principal technical feasibility
- Close coordination with WP5 “Demonstration of the TRUSTS Platform in three business-oriented use cases” to align with commercialization planning in task T7.5
- Close coordination with WP8 aligning on and creating synergies between stakeholder engagement, community building and project communication.

Next Steps

Current activities will be continued. Additionally, the Business & Exploitation Lead, utilizing T7.6 “Innovation Impact Insurance” as primary conduit, will additionally focus on collaboratively developing answers to the following critical questions (selection):

1. How can TRUSTS foster an agreement between the TRUSTS consortium partners, balancing recognition of their IPR and commercial interests of a future operator?
2. How can TRUSTS attract and / or facilitate establishing of an operating company to ascertain operationalization and commercialization of project outputs?

© TRUSTS, 2021

3. How can TRUSTS meaningfully and at scale attract early adopters, particularly data buyers and sellers, as well as data markets for federation, whilst the platform is under development and no operating company is in place, yet?
4. How can TRUSTS optimize the exploitation of R&D outputs around the defined three UCs with respect to data trading as opposed to a sole focus data exchange & privacy-preserving processing?

How does TRUSTS contribute and link to artefacts and participants within the evolving European data economy, leverage and collaborate with parallel national and pan-European initiatives and projects, and ascertain a meaningful and sustainable contribution in support of the European Data Strategy?

5 Data Management Plan

5.1 Overview

Concerning the management of datasets created, processed and published within the TRUSTS project, the Data Management Plan (DMP) (D1.6) provides information on TRUSTS data management policy and key information. This includes the organisational and technical measures regarding data collection, handling and storage of data, as well as key aspects such as the responsibilities of the respective project partners, the compliance with the FAIR data principles (Findable, Accessible, Interoperable, and Reusable) and information on data volume, access, licensing and integration features, in accordance with the relevant legal framework and in particular the GDPR. The DMP will be regularly updated to reflect the development and progress of the project in terms of Data Management. This report contains all updates of the TRUSTS DMP document from July to December 2021 (M19 – M24). For previous changes see Annual Public Report I (M12) and Periodic Technical Report (M18).

5.2 Purpose

The project's DMP lists all relevant information on current and planned data management activities. It is based on the template for the ERC Open Research DMP. In summary, the management of research data in the TRUSTS project is based on the following rules:

- Provide a maximum level of security for sensitive data and personal data, including the exchange of personal and/or sensitive data between selected partners
- Use well-known, established repositories for publishing and archiving non-sensitive research data
- Encourage data providers to make non-sensitive data available using Creative Commons licenses
- Raise awareness among researchers, companies and public stakeholders for the importance of making non-sensitive data available to the public

5.3 State of the art

Data is constantly growing. Unlike other raw materials such as fuels of the economy, e.g. coal or oil, data has the advantage that their supply does not end at some point, but rather they multiply inexorably. Almost regardless of what an individual or a company does, vast amounts of data are created in nearly every

© TRUSTS, 2021

industry process. Whether it's shopping online, using an app, or sending an email every day – information is collected everywhere, providing further information about, e.g. corporate strategy, purchasing behavior or the wishes of the respective person. And this information can be used economically, because unlike oil, data is not consumed through use, but gains its actual value in the process. The social significance of data streams must also be kept in mind. There is a lot of potential in them, especially if they are made technologically usable in a sensible way.

Data has become one of the most important raw materials and is of high importance in nearly every industry sector worldwide.[17]. When it comes to the state of the art, the following three dimensions are most relevant to consider: Digital Single Market (DSM), aiming at fully unleashing the data economy, allowing a free flow of data and therefore enabling companies and public stakeholders to store and process non-personal data wherever they choose in the EU. On the other hand, the GDPR provides the free movement of personal data within the Union, next to its primary goal of protecting personal data. Lastly, the FAIR principles, which were established in 2016 by the FORCE 11 group, focusing on the optimal preparation of research data for humans and machines. In this context, FAIR data does not necessarily mean completely open data, but rather data to be 'as open as possible, as closed as necessary'. This is especially important when dealing with industrial partners and building sustainable business models.

These three dimensions have to be considered by today's (data) platforms. Data is traded, exchanged and published in a trustworthy and secure way providing clear legal and ethical frameworks, where data based services, related software and tools are offered and easily used, and where data professionals can receive training to improve their knowledge and skills. Within the TRUSTS project, datasets of various natures are collected, processed or generated. This includes data that is already existing, e.g. anonymized CRM data, and new data such as metadata and project management data that is created as the project progresses dynamically.

5.4 DSM, GDPR and FAIR data activities by partners

Five out of seventeen project partners have provided updates on data types, size and formats. SWC has added RDF in any serialization as an expected data format. FhG states that they do not process demonstrative data any longer. TUD has changed the data format of the collected documents for creating a taxonomy of data marketplace from textual (.docx) to tabular (.xlsx). RSA and NOVA have added the data size *small (<1GB)* for their processing data. For LUH, KNOW, DIO, LST, EBOS and REL, the status remains the same as in M18.

Three out of seventeen partners provided updates on FAIR and data security. The deliverable D3.7 of SWC is available via the TRUSTS website so that it is findable and accessible. For NOVA data findability is not applicable because their data is internal and will be used anonymously in the TRUSTS trials in compliance to NOVA processes. Data analysis will securely be achieved within NOVA premises using PSI and other applications and data onboarding mechanisms/interfaces deployed by TRUSTS. In addition TRUSTS will provide, according to the deliverable D2.2, the respective interfaces and the data analysis/onboarding process. Last but not least, as also defined in D2.2, TRUSTS will establish the appropriate mechanisms to check the integrity and security of the applications that will be onboarded to the TRUSTS data marketplace. All transactions should be logged appropriately to maintain the appropriate quality, security and traceability levels. The outcome of the analysis will only be visible to NOVA responsible employees and will be used for

[17] Cost-benefit analysis for FAIR research data, 2018, https://op.europa.eu/en/publication-detail/-/publication/d375368c-1a0a-11e9-8d04-01aa75ed71a1

the evaluation of the MVP version provided by TRUSTS. Therefore the data is not publicly accessible beyond the project duration and scope. The data does not need to be interoperable as they will not be shared in the project as described in the respective UCs and the DoA. Nevertheless, should it be necessary for dissemination and exploitation purposes NOVA could consider creating sample (fake) data adequate to demonstrate the functionality of the TRUSTS datamarket place for extra-TRUSTS stakeholders. This data is then made reusable.

PB provides internal data which will be used strictly in the scope of the TRUSTS trials according to the organization's regulations and processes. Therefore PB data is not publicly FAIR. The data will be anonymized, aggregated and be used within the premises of PB as per the description of UC2. For data analysis the TRUSTS mechanisms and services will be used (e.g., PCI, MPC, onboarding services etc.). Analysis results will only be visible to PB so that the evaluation process defined in D2.2 will be conducted. For this the security and protection preservation mechanisms that will be provided in the context of the project (WP4) are used. The PB data will be available only within the project's scope and duration. This data does not belong to PB and will not be licensed for any kind of distribution. Therefore the data will also be used according to Greek and European laws and regulations.

EBOS added that EBOS and UC1 data is used in the scope of the project and WP5 trials (supporting the UC1 trials) and will not be available beyond the project duration. For LUH, KUL, RSA, DIO and FORTH the status remains the same as in M18.

The DMP update of the remaining project partners, which are not mentioned in any of the Reports in M12 or M18, on DSM, GDPR and FAIR data activities is still pending or they do not work with data, so they cannot (yet) make any statements. Changes will be addressed in the next update in M36.

5.5 Processed and published data(sets) as of December 2021

Beside the published deliverables with included data on the TRUSTS website three out of 17 partners have published their datasets elsewhere on the web. KNOW has specified their first published dataset in Zenodo which is used for studying privacy aspects of recommender systems. More datasets for this research are planned to be shared in the last project year of TRUSTS.

TUD has reported three publications on the repository *4TU.ResearchData*, one in collaboration with IDSA. The first one is a list of academic articles to investigate the state of the art of data marketplace research. This data is also part of the deliverable D2.1 *Definition and analysis of the EU and worldwide data market trends and industrial needs for growth*. The second dataset contains collected documents for creating a taxonomy of data marketplaces, which is a supplementary document of the article *Creating a Taxonomy of Business Models for Data Marketplace*. The third one was created in collaboration with IDSA and is an overview of literature and other sources used for market research (e.g. journal articles and conference papers). It is part of the article *Why open government data initiatives fail to achieve their objectives: Categorizing and prioritizing barriers through a global survey*. All other participants have not published any further data. This will be done in M36, if applicable.

The datasets mentioned are all publicly and openly accessible in a repository. They are all provided with a DOI and are licensed CC BY 4.0. The metadata can be exported to different formats such as BibTeX, DataCite or Dublin Core, as well as in other formats depending on the repository.

© TRUSTS, 2021

5.6 Next steps

As new data types and workflows are likely to emerge during the project duration, which still occur in the DMP, updates of the DMP will be provided as per the DoA in M36, which will show in-depth information on the processed datasets and workflows.

6 Conclusions and Next Actions

Looking back at the second project year 2021 (M13-M24), it can be summarized that the TRUSTS project proceeded as planned.

In terms of WP1, aside from the many day to day project management tasks, the consortium was adequately prepared for the midterm review and periodic reporting. The midterm review itself was organised and executed successfully in close alignment with the EC. Among the implementation of the valuable feedback received at the midterm review, the project management will proceed with the quality control and submission of the remaining 22 (out of 70) deliverables. In the scope of WP2 the analysis of worldwide data marketplace ecosystems was initiated, the second iteration and finalisation of the Functional Requirements (FR) was set up, the evolution of the trials evaluation testing methodology and the business evaluation methodology was updated and the definition of the final version of the TRUSTS platform architecture was achieved. While WP2 ends at M24, the results will continue to inform the UCs, business plans for the UCs, technical decisions, strategic decisions as well as the project as a whole. WP3 has set up and managed the development environment, supported partners in terms of resources, roles, access, deployment instructions and documentation and provided hands on support, using CI/CD pipelines (using Jenkins) and gitlab repository to ease the development. Additionally, WP3 started to set up a blockchain demonstrator for smart contracts, further developed plans for interoperability with data markets and EOSC initiatives, generated the TRUSTS Knowledge Graph, developed MVPv1 based on the IDS Trusted Connector and the CKAN framework and refined six recommendation use cases. A clear pathway and next steps in terms of infrastructure, smart contracts, data marketplace interoperability solutions, data governance, platform development and user/corporate profile brokerage is defined and layed out in this deliverable. In terms of WP4, Cryptographic primitives involved in building collaborative trust systems were investigated including Fully homomorphic encryption (FHE) and Secure multi-party computation (MPC). A private, efficient, and secure TL method (CryptoTL) was proposed and tested, six risk analysis modules were developed, federated learning methodologies were evaluated and analysed. Further work will explore VFL methods in terms of collaborative analytics and preserving data privacy, finalising research regarding SHAP values as input for VFL, continuation of the implementation of prototypes and enhancing the developed anonymization algorithm. WP5 mainly focused on setting up the test environment as well as the relevant planning and operational management for the execution of the first phase of the three UC trials. The first report laying out the overall plan of the first cycle of the TRUSTS trials was produced as well. The second cycle of the TRUSTS UC trials are set to commence in January 2022, considering some adaptations based on the feedback of the midterm review. The major outcomes of WP6 include an overview of legal frameworks for the project to comply with the principles of research ethics; analysis of European laws and regulations relevant to data transactions and the TRUSTS Platform development; definition of legal and ethical requirements and identification of potential legal and ethical obstacles and recommendations for policy makers and stakeholders. Further work will include research on new legal frameworks discussed in the EU such as the Data Governance Act, Digital Services Act and Digital Markets Act as well as further development of legal and

© TRUSTS, 2021

ethical requirements and recommendations for relevant stakeholders and policy makers on identified legal gaps and lessons learned. Some of the major outcomes of WP7 include a business-model centric and unified taxonomy that forms the basis to choose among viable business model options, a drafted approach to exploitation and commercialization, a taxonomy to inform auxiliary services as well as alignment among all WPs in terms of the innovation impact assurance. WP8 performed the day to day work as per the communication and dissemination strategy, visual identity and promotion activities were extended and a podcast section was added and collaborations with other projects were established. Further steps will include collaborative work with our SAB, engaging them with the European Data community. Events are planned for Q1 and Q4 of 2022 to consolidate stakeholder engagement. Additionally, the training and capacity building programme will be implemented. The objectives of WP9 have mostly been achieved within the first year of the project. One major outcome of the second year was that pending ethical concerns were addressed in terms of D9.9 in March 2021, which consolidated pending concerns of D9.1, D9.4, D9.5, D9.8 and D9.9. Ethics compliance is a constant follow up work which requires active engagement by all partners and will be closely monitored until the end of the project.

This work is further solidified by our Scientific Lead, Technical Lead, Innovation Lead, Security Lead, Legal & Ethical Lead, Communication and Community Lead as well as our Business and Exploitation Lead. Within the second project year, TRUSTS increased its participation in the scientific discourse on topics such as trusted and secure data sharing platforms, privacy-aware analytics methods, ICT and Data skills, challenges for IT standardisation and more. Under the management of the Scientific Lead, the project increased its number of publications by seven, bringing the overall number of published research papers to nine. Ten more publications are planned for the third project year. The Technical Lead established consensus amongst technical project partners with regards to strategic questions in terms of architectural paradigm, the reuse of existing software and non-FR of strategic importance. Next steps include further iteration on architecture specification, if needed and the further refinement of the architecture. The Innovation Lead implemented the innovation process, maintained the innovation registry and closely monitored and reported the innovation activities in terms of the Innovation Radar. Further steps include the finalisation of business and remuneration models and the further promotion of activities to increase commercialisation potential. The Security Lead advanced the project's understanding of security issues and potential vulnerabilities. Next steps include the further implementation of security solutions especially with regard to GDPR, ethics and regulatory issues. The Legal & Ethical Lead provided an overview of legal frameworks, identified potential obstacles and provided recommendations to overcome them. Next steps include further research and analysis of legal frameworks discussed in the EU such as the Data Governance Act, Digital Services Act and Digital Markets Act. The Communication and Community Lead guided and expanded the outreach, focusing the communication efforts more on the project's progress and achievements. Strengthening ties with networks and programmes such as Safe-DEED, DOME4.0 or i3-market and engaging the SAB with the European Data community. The Business & Exploitation lead guided and coordinated the alignment among partners on important topics as per WP7 e.g. Lessons learned from DMA, Platform federation, TRUSTS business model, and the Datamarket platform operator, etc. Further work will focus on key questions pertaining the balance of IPR and commercial interests, the potential operation of an entity commercialising TRUSTS, attracting early adopters and more.

© TRUSTS, 2021