



## TRUSTS Trusted Secure Data Sharing Space

### **D2.2 Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition I**

#### **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	<a href="https://trusts-data.eu/">https://trusts-data.eu/</a>		
Deliverable	D2.2 Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition I		
Work Package	WP2		
Contractual due date	30/06/2020	Actual submission date	26/6/2020
Nature	Report	Dissemination Level	Public
Lead Beneficiary	FNET		
Responsible Author	Ioannis Markopoulos		
Contributions from	SWC, PA, eBOS, LST, REL, FORTH		



**Revision history (including peer reviewing & quality control)**

Version	Issue Date	% Complete	Changes	Contributor(s)
v0.1	10/02/2020	5	Initial Deliverable Structure	Ioannis Markopoulos (FNET)
v0.2	31/3/2020	20	Initial input to UCs, background information and interviews	Ioannis Markopoulos (FNET), George Margetis (FORTH), George Kostopoulos (PB), Aspa Filippou (eBOS), Christos Skoufis (eBOS), Christos Roupas (REL), Ana Maria Florea (REL), Thomas Thurner (SWC)
v0.3	05/05/2020	20	Additional section included	Ioannis Markopoulos (FNET)
v0.4	15/05/2020	80	Requirements analysis	Ioannis Markopoulos (FNET), George Margetis (FORTH), Stavroula Ntoa (FORTH)
v0.5	22/05/2020	80	Requirements analysis	Ioannis Markopoulos (FNET), George Margetis (FORTH), Stavroula Ntoa (FORTH)
v0.6	23/05/2020	80	Requirements analysis	Ioannis Markopoulos (FNET), George Margetis (FORTH), Stavroula Ntoa (FORTH)
v0.7	24/05/2020	90	Functional Requirements definition	Ioannis Markopoulos (FNET), George Margetis (FORTH), Stavroula Ntoa (FORTH)
v0.8	25/05/2020	95	Statistical analysis	George Margetis (FORTH), Stavroula Ntoa (FORTH), Mark de Reuver (TUD), A.E.Abbas (TUD)
v0.9	26/05/2020	98	Functional Requirements definition (revised)	Ioannis Markopoulos (FNET), George Margetis (FORTH), Stavroula Ntoa (FORTH)
V1.0	27/05/2020	100	Deliverable ready for peer review	Ioannis Markopoulos (FNET)
V1.1	09/06/2020	100	Peer review comments accommodation	Ioannis Markopoulos (FNET), George Margetis (FORTH)
V1.2	17/06/2020	100	Comments accommodation	Ioannis Markopoulos (FNET), George Margetis (FORTH)
V1.3	19/06/2020	100	Comments accommodation & final editing	Ioannis Markopoulos (FNET)

## 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 authors(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 in negligence or otherwise howsoever 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

1	Introduction.....	11
1.1	Mapping Projects' Outputs .....	12
1.2	Deliverable Overview and Report Structure .....	13
2	Methodology .....	14
3	TRUSTS Objectives Context for the Requirements Elicitation & Analysis .....	15
4	Current data marketplace initiatives and industry needs .....	16
4.1	Qualitative Requirements Elicitation per Industry Segment .....	17
4.1.1	Industry 4.0 .....	17
4.2	Requirements.....	19
5	Related legal and regulatory framework - Platforms, Free Flow of Data and Data Market Place .....	20
5.1	General Data Protection Regulation.....	21
5.2	e-Privacy Regulation Proposal.....	22
5.3	Free Flow of Non-Personal Data Regulation .....	23
5.4	Platform-to-Business .....	24
6	Questionnaires and Interviews.....	25
7	Stakeholders feedback analysis for a commercial financial and operators' industry vertical data marketplace platform .....	26
7.1	Electronic Survey .....	26
7.1.1	Method and Procedure .....	26
7.1.2	Participants.....	27
7.1.3	Results.....	30
7.1.4	Requirements .....	44
7.2	Interviews .....	53
7.2.1	Method and Procedure .....	53
7.2.2	Participants.....	53
7.2.3	Requirements .....	55
8	Detailed Use Case Definition and Target Business & Technical KPIs .....	58
8.1	Use Case 1 - The (AML) compliance use case .....	58
8.1.1	UC1 data marketplace requirements and trial definition .....	59
8.1.2	UC1 roles .....	62
8.1.3	UC1 KPIs .....	63
8.1.4	UC1 longer terms business KPIs .....	65
8.2	Use Case 2 - Agile Marketing through Data Correlation.....	66
8.2.1	UC2 data marketplace requirements and trial definition .....	67
8.2.2	UC2 roles .....	68
8.2.3	UC2 KPIs .....	69
8.2.4	UC2 longer term business KPIs.....	70
8.3	Use Case 3 - The data acquisition to improve customer support services use case .....	71
8.3.1	UC3 data marketplace requirements and trials definition .....	73

8.3.2	UC3 roles .....	75
8.3.3	UC3 KPIs .....	76
8.3.4	UC3 longer terms business KPIs .....	77
8.4	UC requirements .....	79
9	TRUSTS Requirements Analysis and Categorization .....	81
10	Conclusions and Next Actions.....	86
	Annex I: Questionnaire .....	88
	Annex II: Interviews guidelines.....	97
	Annex III: Interviews .....	98
	Annex IV: DMA Insights.....	115

## List of Figures

Figure 1: Methodology for the requirements elicitation, analysis and usage .....	15
Figure 2: Participating organizations' information .....	27
Figure 3: Roles of participants' organizations in the data exchange value chain.....	28
Figure 4: Participants' roles and level of management in their organizations.....	28
Figure 5: Participants' business experience .....	29
Figure 6: Participants understanding and involvement in data sales / buying process in their organization .....	29
Figure 7: Data buying frequency and type .....	30
Figure 8: Data selling frequency and type.....	31
Figure 9: Frequency and type of provided data applications or services .....	32
Figure 10: Frequency and type of data applications or services bought .....	33
Figure 11: Data marketplace functions.....	34
Figure 12: Anonymization (desired functionality) – percentages of responses within organization roles' categories.....	35
Figure 13: De-anonymization risk analysis (desired functionality) – percentages of responses within organization roles' categories .....	36
Figure 14: Datasets catalogue (desired functionality) – percentages of responses within organization roles' categories.....	36
Figure 15: Datasets valuation (desired functionality) – percentages of responses within organization roles' categories.....	37
Figure 16: Datasets search discovery (desired functionality) – percentages of responses within organization roles' categories .....	37
Figure 17: Standardization information (desired functionality) – percentages of responses within organization roles' categories .....	38
Figure 18: Data marketplace functions: desired vs. required .....	39
Figure 19: De-anonymization risk analysis (mandatory) – percentages of responses within organization roles' categories.....	40
Figure 20: Multiparty computation (mandatory) – percentages of responses within organization roles' categories.....	40
Figure 21: User authentication (mandatory) – percentages of responses within organization roles' categories .....	41
Figure 22: Dataset valuation (mandatory) – percentages of responses within organization roles' categories.....	41
Figure 23: Dataset rating (mandatory) – percentages of responses within organization roles' categories .....	42
Figure 24: Multiparty computation (mandatory) – percentages of responses within organization roles' categories.....	42
Figure 25: Pricing models.....	43
Figure 26: Pay per use – percentages of responses within organization roles' categories .....	44
Figure 27: The AML process through the TRUSTS data marketplace .....	59
Figure 28: UC1 Partners Roles.....	63
Figure 29: UC2 agile marketing trial architecture .....	66
Figure 30: High-level conceptual architecture of the “Agile marketing activities through correlation of anonymized telecom and banking” use case .....	69

Figure 31: UC3 Structure ..... 74

Figure 32: UC3 Actors ..... 76

## List of Tables

Table 1: Adherence to TRUSTS GA Deliverable & Tasks Descriptions .....	12
Table 2: Interviewees consolidated information.....	53
Table 3: UC1 high level trial scenarios .....	61
Table 4: UC1 performance and process KPIs.....	63
Table 5: UC1 longer term business KPIs.....	65
Table 6: UC2 high level trial scenarios .....	68
Table 7: UC2 performance and process KPIs.....	69
Table 8: UC2 longer term business KPIs.....	71
Table 9: UC3 high level trial scenarios .....	74
Table 10: UC3 performance and process KPIs.....	76
Table 11: UC3 longer term business KPIs.....	78
Table 12: TRUSTS Functional Requirement Specifications.....	81



## Glossary of terms and abbreviations used

Abbreviation / Term	Description
AI	Artificial Intelligence
AML	Anti-Money Laundering
E2E	End-to-End
DaaS	Data as a Service
DoA	Description Of the Action
FAIR	Findable, Accessible, Interoperable, and Reusable
FFNPDR	Free Flow of Non-Personal Data Regulation
FR	Functional Requirements
GA	Grant Agreement
GDPR	General Data Protection Regulation
IR	Interview Requirements
KPI	Key Performance Indicator
KYC	Know You Customer
LR	Legal/Regulatory Requirements
ML	Machine Learning
MPC	Multi Party Computation
NPS	Net Promoters Score
P2BR	Platform-to-Business Regulation
PEP	Politically Exposed Person
PoV	Point of View
PSI	Private Set Intersection
QR	Questionnaire Requirements
SLA	Service Level Agreement
SR	Industry Requirements
SUS	System Usability Scale
UC	Use Case
UCD	User Centred Design
UCR	Use Case Requirements

## Executive Summary

The purpose of this deliverable is to report the work performed in the context of Task 2.2: Industry-specific functional requirements elicitation and analysis (telecom, financial, corporate/personal data). The main analysis of this deliverable is related to:

- a) compilation of current data marketplace initiatives, industry related needs, features and capabilities as well as regulatory trends and legislation
- b) requirements analysis and E2E service definition
- c) establishment of the targeted data marketplace functions for the financial and operators' sector and the vertical and cross functional use cases aiming at demonstrating and benchmarking the E2E data marketplace operation and value added to the industry

In order to collect and analyse the requirements we worked in the following axis:

- Questionnaires targeting all actors in the data marketplace ecosystem
- Interviews with respective industry representatives
- Collection and analysis of requirements from previous respective initiatives
- Requirements from the use case participants representing the telecom and financial sectors.

Lastly, the TRUSTS objectives were analysed in terms of their impact on the requirements.

The final expected outcome of TRUSTS will be the successful implementation of a federated data marketplace which will be evaluated by three specifically designed use cases, carefully extracted from real-life scenarios.

These use cases are:

- UC1: Smart big-data sharing and analytics for Anti-Money Laundering (AML) compliance.
- UC2: Agile marketing activities through correlation of anonymized banking and operators' data.
- UC3: Buying data from a data marketplace to improve Natural Interaction.

The requirements analysis from the sources, as described above, provides actual representation of the desired functionalities of the TRUSTS platform e.g. datasets and services onboarding to the TRUSTS data marketplace, subscribers' management, data governance and protection, etc.

It also includes the definition of high level UC scenarios<sup>1</sup> aiming at demonstrating the potential of the TRUSTS technology and all of its innovative features.

This deliverable constitutes the first of the two reports, foreseen to be produced in the context of the project, containing the detailed analysis of the requirements for a commercial financial and operators' industry vertical data marketplace platform and the use cases definition including the target KPIs that would set the benchmarking for the actual measurements according to the Task 2.3 methodology.

---

<sup>1</sup> Please note that the UC scenarios and test cases will be refined in the subsequent project phases using the Task 2.3 methodology which will be available at M6 of the project development.

# 1 Introduction

The lack of trusted and secure platforms and privacy-aware analytics methods for secure sharing of personal data and proprietary/commercial/industrial data hampers the creation of a data market and data economy by limiting data sharing mostly to open data. This trend will continue unless different technical standards, quality levels, and legal aspects are allowed to diverge uncontrollably.

**TRUSTS will ensure trust** in the concept of data markets as a whole via its focus on developing a platform based on the experience of two large national projects, while allowing the integration and adoption of future platforms. The TRUSTS platform will act independently and as a platform federator taking into account the legal and ethical aspects that apply on the entire data valorisation chain, from data providers to consumers. To that end, it will (a) Set up a fully operational and GDPR-compliant European Data Marketplace for personal related data and non-personal related data targeting both personal and industrial use by leveraging existing data marketplaces (International Data Space and Data Market Austria) and enriching them with new functionalities and services. (b) Demonstrate and realize the potential of the TRUSTS Platform in 3 use cases targeting the industry sectors of corporate business data, specifically in the financial and telecom operator industries while ensuring it is supported by viable, compliant and impactful governance, legal and business models.

To create a European Data Market based on secure and trustworthy data exchanges, the TRUSTS consortium brings together technology providers that are already deeply involved in major national data market projects.

This deliverable constitutes the first version of the two reports containing the detailed analysis of the requirements for a commercial financial and operators' industry vertical data marketplace platform and the use cases definition including the target KPIs that would set the benchmarking for the actual measurements.

## 1.1 Mapping Projects' Outputs

The purpose of this section is to map TRUSTS Grand Agreement commitments, both within the formal deliverable and task description, against the project's respective outputs and work performed.

Table 1: Adherence to TRUSTS GA Deliverable & Tasks Descriptions

TRUSTS Task		Respective Document Section(s)	Justification
T2.2 Industry-specific functional requirements elicitation and analysis (telecom, financial, corporate/personal data)	<i>The aim of this task is to capture requirements from the financial institutions, telecom operators and corporate data providers/users, as well as from industrial and regulatory stakeholders. It will include the articulation of detailed use case scenarios and usability needs, and relevant target technological and business KPIs which will be validated in the pilots. Specifically, this task will be divided into three key sub tasks:</i>	Section 5	<i>In Section 5 an analysis of selective current data marketplaces initiatives are analysed based on respective DMA data collection.</i>
	<i>(a) systematic compilation of current data marketplace initiatives, industry related needs, features and capabilities as well as regulatory trends, legislation and standardisation,</i>	Section 6	<i>In Section 6 regulatory and legislation needs are analysed based on respective Safe-DEED work.</i>
	<i>(b) requirements analysis and E2E service definition,</i>	Section 7	<i>In section 7 the questionnaires and the interviews with respective stakeholders including telecommunication and banking sectors are referred.</i>
	<i>(c) establishment of the targeted data marketplace functions for the financial and operators' sector and the vertical and cross functional use cases aiming at demonstrating and benchmarking the E2E data marketplace operation and value added to the industry. The requirements' capture will involve the documentation of industrial and regulatory needs and opinions about new innovative data</i>	Section 8	<i>In Section 8 the individual requirements are analysed in a way that constitute a comprehensive set of functional requirements in order to drive platform implementation as well as to be used through the T1.3 methodology for the evaluation of the TRUSTS environment.</i>
		Section 9	<i>In Section 9 the UCs are thoroughly analysed for their needs with respect to the TRUSTS platforms as well as the high level trials descriptions and respective KPIs.</i>

	<i>marketplace service vertical, which will set the baseline for conducting the actual measurements during the use case trials. This task will be performed via electronic surveys, prepared by FNET, and through dedicated workshops with the foreseen actors of every case study, animated by FNET and hosted by every case study owner.</i>	<i>Section 10</i>	<i>In Section 10 there is a comprehensive production of the Functional Requirements (FRs) with reference to the source of each requirement and the Task that will undertake the implementation.</i>
<b>TRUSTS Deliverable</b>			
<p><i>D2.2: Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition I</i></p> <p><i>First version of the two reports containing the detailed analysis of the requirements for a commercial financial and operators' industry vertical data marketplace platform and the use cases definition including the target KPIs that would set the benchmarking for the actual measurements.</i></p>			

## 1.2 Deliverable Overview and Report Structure

This deliverable adopts a comprehensive approach for the definition of the functional requirements starting from the methodology and then proceeding to the elicitation of requirements per source, their analysis, and the definition of the functional requirements towards the TRUSTS platform implementation WPs.

In particular, in Section 3 the Methodology is analysed and the requirement sources are defined. Section 4 provides the context of the TRUSTS objectives in which the requirements are collected and analysed. Section 5 summarizes the industrial needs and requirements. Section 6 analyses the respective Legal and Regulatory framework as well as the respective requirements. In Section 7 the questionnaires and interview methodology are presented while in Section 8 the responses are qualitatively and quantitatively analysed, and the respective requirements are depicted. In Section 9 the UCs are described and their requirements, high level scenarios and KPIs are identified. In section 10 the complete set of the functional requirements (FRs) is produced with reference to the respective source requirement and the Task within the project that will undertake the respective implementation.

## 2 Methodology

The TRUSTS project follows a user centred approach, which places the stakeholders of a product or a system at the centre of its design and development. User centred design (UCD) seeks to answer questions about users and their tasks and goals, such as “who are the users of the TRUSTS platform?”, “what are user tasks and goals?”, “what expectations do users have from this platform?”. The answers to these questions are then used to drive the design and development. This is achieved by involving and talking directly to key stakeholders throughout the project, starting from its very beginning, in order to assure that the platform will deliver the foreseen requirements.

Following the Ergonomics of Human System Interaction standard (ISO 9241-210<sup>2</sup>), which is part of the multi-part standard ISO 9241 and a revision of the withdrawn ISO 13407:1999, outlines four essential activities in a user-centred design project:

1. Requirements gathering - Understanding and specifying the context of use;
2. Requirements specification - Specifying the user requirements;
3. Design - Producing design solutions;
4. Evaluation - Carrying out user-based assessment of the TRUSTS platform.

Task 2.2 focuses on the two first activities presented above which deal with the collection, analysis and specification of the requirements.

The means to collect the requirements are through:

1. key stakeholders’ interviews
2. the use of dedicated electronic surveys
3. the analysis of selective related data marketplace activities
4. the analysis of related legal framework
5. the analysis of the use cases

The context, in which the requirements are analysed, is the TRUSTS project objectives.

The methodology to produce and use the TRUSTS data marketplace FRs is illustrated in Figure 1. In particular:

- All requirements sources are analysed for individual requirements and their justification
- The abovementioned requirements drive the definition of the FRs, which will be used for the implementation of the TRUSTS platform as well as the operational processes design.
- To assist implementation, each FR is mapped to the respective project task.
- Furthermore the FRs will be used by the methodology defined in task T2.3 entitled “Testing framework and benchmarking” in order to evaluate and provide coherent feedback through the UC trials (WP5).

---

<sup>2</sup> <https://www.iso.org/standard/77520.html>

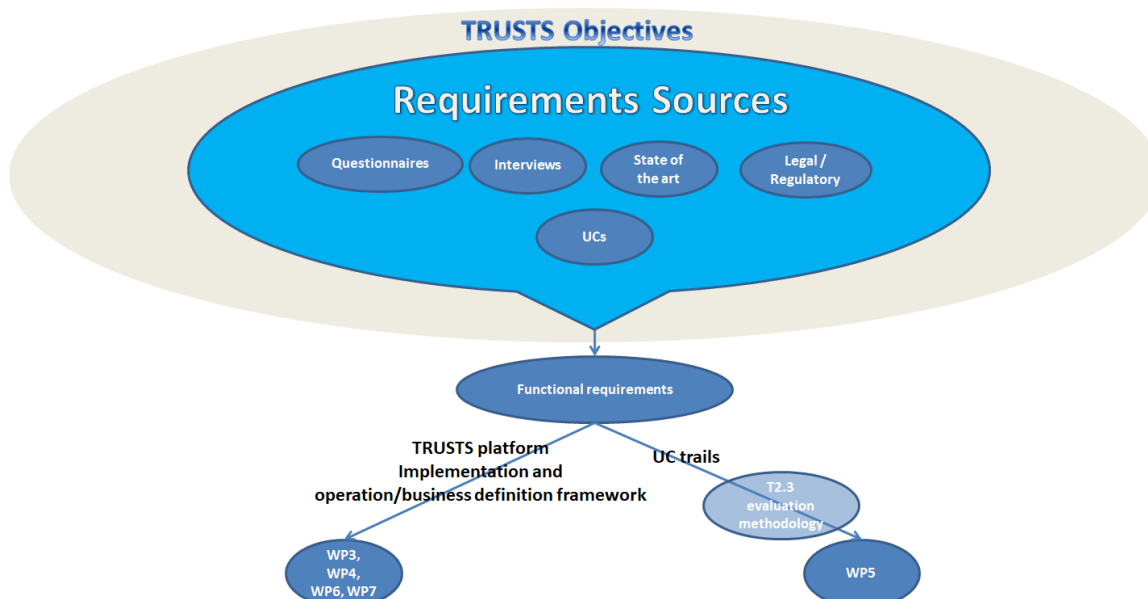


Figure 1: Methodology for the requirements elicitation, analysis and usage

This process is depicted in the implementation of deliverable 2.2: Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition I (first version, M6), as well as in deliverable 2.3: Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II (Final version, M24), which will essentially describe an updated version of the requirements.

### 3 TRUSTS Objectives Context for the Requirements Elicitation & Analysis

The main objective of the TRUSTS project is summarized as follows:

TRUSTS will ensure trust in the concept of data markets as a whole via its focus on developing a platform based on the experience of two large national projects, while allowing the integration and adoption of future platforms by means of interoperability. The TRUSTS platform will act independently and as a platform federator, while investigating the legal and ethical aspects that apply on the entire data valorisation chain, from data providers to consumers, i.e., it will:

1. setup a fully operational and GDPR-compliant European Data Marketplace targeting individual and industrial use by leveraging existing data marketplaces (Industrial Data Space, Data Market Austria) and enriching them with new functionalities and services to scale out.
2. demonstrate and realize the potential of the TRUSTS Platform in 3 use cases 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.

**Remarks:**

- In the requirements elicitation process both IDSA and DMA representatives provided requirements through interviews and background information<sup>3</sup>.
- The legal framework for personal and non-personal data exchange is analysed
- All the three use cases provided their needs for the required TRUSTS environment services, performance, KPIs and high-level scenarios.

Relevant to this deliverable is TRUSTS part of objective 1:

Objective 1	WP 2 Requirements Elicitation & Specification
To analyse the EU & worldwide challenges and trends for <b>data-sharing</b> and define the requirements for the provision of a <b>multi, concurrent and cross-domain, secure and scalable end-to-end (E2E) data marketplace service</b> .	
Achieving this objective will require capturing and eliciting <b>end-user requirements</b> , as well as a detailed analysis of end-user needs in view of transforming these into specific functional requirements and an architectural design.	
<b>Measurable outcomes:</b> <ol style="list-style-type: none"> <li>1. Detailed <b>industry-specific functional specifications</b> appropriate for a data marketplace linked to specific target <b>KPIs</b> considering and bridging the vertical user point of view (PoV) with the analytics/solution provider PoV and the data marketplace platform provider PoV produced by M6;</li> </ol>	

**Remarks:**

- Relevant industrial stakeholders provided their PoV through both questionnaires and interviews. In addition, existing DMA consultation with stakeholders were used to assist the requirements elicitation process. Use Cases participants provided their requirements for services, performance and respective KPIs. Analysis of the elicited requirements led to the production on the first version of the TRUSTS platform functional requirements and specifications.

## 4 Current data marketplace initiatives and industry needs

This section aims at analysing the state of the art in the data marketplace domain to acquire an insight on the respective requirements and best practices that may be adopted by TRUSTS.

Such analysis regarding best business practices and state-of-the-art technologies is performed for the completeness of the requirements elicitation endeavour, no matter if the project is the early stages of its implementation. Further analysis and trends in the data marketplace ecosystem will be described in the deliverables D2.1 and D7.1 on M18 of the project.

<sup>3</sup> Please note that the D2.2 deliverable focuses on producing the TRUSTS functional requirements. Parallel tasks defining the TRUSTS architecture capitalizing on IDS and DMA developments are progressing within WP3.



TRUSTS is initially aiming at offering services to the financial and telecommunication domains. Nevertheless, due to the fact that

- businesses are not vertical silos anymore but they aim expanding to other domains and
- in the TRUSTS exploitation analysis we will aim at offering sustainable data marketplace services to a wide range of business sectors beyond the telecom and financial ones, we elicit requirements from key industrial sectors.

TRUSTS aim is to implement a platform, which can be incorporated to the value chain of various industries aiming at effectively and efficiently collaborating in order to innovate and provide value added services.

Such information was retrieved building upon respective analysis achieved in the DMA project.

## 4.1 Qualitative Requirements Elicitation per Industry Segment

Data for this chapter coming from a workshop series carried out by the Data Market Austria in 2017 with representatives of various industry branches. We focus on Industry 4.0 requirements, which are more relevant to the TRUSTS project<sup>4</sup>. Nevertheless, other domains have been investigated as well i.e. Earth Observation, Mobility, and Energy. Such data will be evaluated towards defining the TRUSTS business model and exploitation actions.

### 4.1.1 Industry 4.0

Even though Industry 4.0 is strongly connected to data centric approaches, experts observe that current activities often still focus primarily on process optimization rather than on integrations into a data centric product value chain.

As a horizontal approach to Industry 4.0 seems to be the key for an intensified need regarding data and service exchange, relevance of the Data Marketplace to the Industry 4.0 domain may be called medium.

A summary of the respective requirements follows:

#### **Data and data services required**

##### Data and services

- All relevant Open Data as a basis
- Basic data for Industry 4.0 (Machines, Sensors, etc.)
- Basic generic services
- Matchmaking services which act as bridge builder for other domains, in recruiting, to research
- Data aggregation / anonymizing services, which allow the processing of personalized data. Such services may be run by a special trustee, who guarantees conformity with privacy regulations.

##### Motivation

- Use Data Marketplace as a vehicle to push Standards for Industry 4.0
- A strong metadata layer as a basis for quality and interoperability
- Transaction based billing models

---

<sup>4</sup> Extensive DMA insights are included in Annex IV.

### Conditions to trade data and services

- Guaranteed quality and legal certainty for the trade
- Standardized SLAs
- A mechanism that ensures that only “serious vendors” trade on the Data Marketplace
- Support in acquiring and marketing by the Data Marketplace, possibly by clustering offers.

### **Technical requirements and features**

Relevance of Volume, Velocity, Variety and Veracity for the organization

- Storing and processing within the cloud is key in many use cases in Industry 4.0, to meet the above listed requirements (e.g.: Data as a Services - DaaS)
- Doing predictive maintenance of machine parks needs a powerful computing infrastructure
- The volume of data which is stored only on possible demand may cause substantial amounts (e.g.: live production data). How to tackle the problem, that also “unused” and so unexploited data is stored.

Technologies used and planned to be used in the future

- There is a high interest in blockchain technology coming from Industry 4.0.

Requirements on function and usability

- It should be possible to browse and explore data and services openly, but when it comes to the commercial case, data and services may have to be exchanged point-to-point. Additionally, this data should also be protected from unauthorized duplication.
- For some branches long-term-preservation of product (-performance) data is key to ensure quality, meet regulations and show long-term stability. In case the Data Marketplace deals with such data, standards for long-term-preservation have to be met and guaranteed.
- User interface has to follow business logic on the first level (browse) and then be detailed for data experts.

Room for experiments and off-trade activity

- Rooms<sup>5</sup> for developing and testing data driven business ideas are important. Often such rooms will be the only place where all facilities (cloud, data, processing, IP, etc.) are present to get data products developed.

### **Basic conditions, legal and policy issues**

Surrounding conditions for the success

- Operational commencement should be done with core functions and core data and services instead of a very broad approach will ensure focus and comprehension.
- The mission as a commercial project should be made clear from the very beginning. The idea of a commercial marketplace should be put in front of the appearance.

### **Conditions for trust**

- Clear statements on the business model and ownership of Data Marketplace convey trust into the superstructure.

---

<sup>5</sup> e.g. Sandboxes that are virtual spaces in which new or untested software or coding can run securely.

- The trust in the data and services provided may be strengthened by a certification process of providers and/or concrete dataset or service provision.

## 4.2 Requirements

The requirements elicited through the above-mentioned review of the current data marketplace initiatives are:

### SR1 Data required by the industry

#### Remarks

Data required:

- All relevant Open Data as a basis
- All types of data may be supported e.g. streams, static data, etc.

### SR2 Data usage processes and services

#### Remarks

- DaaS.
- It should be possible to browse and explore data and services openly, but when it comes to the commercial case, data and services may have to be exchanged point-to-point. Additionally, this data should also be protected from unauthorized duplication.
- Sometimes long-term-preservation of product (-performance) data is key to ensure quality, meet regulations and show long-term stability. In case the Data Marketplace deals with such data, standards for long-term-preservation have to be met and guaranteed.
- User interface has to follow business logic on the first level (browse) and then be detailed for data experts.
- Rooms for developing and testing data driven business ideas are important. Often such rooms will be the only place where all facilities (indicatively cloud, data, processing, IP, algorithm development, etc.) are present to get data products developed.

### SR3 Services required by the industry

#### Remarks

Services required:

- Matchmaking services which act as bridge builder for other domains, in recruiting, to research.
- Data aggregation / anonymizing services, which allow to process personalized data. Such a service. may be run by a special trustee, who guarantees conformity with privacy regulations.
- A strong metadata layer as a basis for quality and interoperability.
- Blockchain technology can constitute a prominent candidate for ensuring information transactions integrity.

**SR4 Operational quality required by the industry.****Remarks**

Operational processes required:

- Transaction based billing models.
- Guaranteed quality and legal certainty for the trade.
- Standardized SLAs.
- A mechanism that ensures that only “serious vendors” trade on the Data Marketplace.
- The trust in the data and services provided may be strengthened by a certification process of providers and/or concrete dataset or service provision.
- Support in acquiring and marketing by the Data Marketplace.

**SR5 Business model requirements****Remarks**

- Operational commencement should be done with core functions and core data and services instead of a “very broad and potentially unclear approach” in order to ensure focus and comprehension.
- The mission as a commercial project should be made clear from the very beginning. The idea of a commercial marketplace should be placed in front of the appearance.
- Clear statements on the business model and ownership of Data Marketplace convey trust into the superstructure.

## 5 Related legal and regulatory framework - Platforms, Free Flow of Data and Data Market Place

TRUSTS aims at analysing the legal and regulatory framework for respective requirements to:

- develop a platform and the respective operational procedures compliant to all respective data privacy and free flow regulations aiming at commercial exploitation of the service
- design trials safeguarding that all involved stakeholders will perform according to the regulations that govern data marketplace operations and data exchange

TRUSTS consortium has concluded to the strategic decision of capitalizing on the respective significant analysis of pertinent HORIZON 2020 projects and especially the project Safe-DEED<sup>6</sup>. This analysis will be used to extract legal and regulatory requirements for the development and operation of the TRUSTS data marketplace. Further analysis will be done at later stages of the project following the trials’ outcome evaluation, which will be

<sup>6</sup> HORIZON 2020 Safe-DEED project (Grant Agreement no 825225), deliverable D3.1 entitled “Legal Frameworks and Ethical Issues”, The deliverable was produced by CiTiP KU LUVEN, Authors: Aleksandra Kuczerawy, Amandine Léonard, Alessandro Bruni, 31/05/2019, [https://safe-deed.eu/wp-content/uploads/2019/07/Safe-DEED\\_D3\\_1.pdf](https://safe-deed.eu/wp-content/uploads/2019/07/Safe-DEED_D3_1.pdf)

included in the deliverable D2.3 entitled “Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II”.

Both personal and non-personal data regulations and directives are taken into account<sup>7</sup>.

## 5.1 General Data Protection Regulation

The Regulation 2016/679 (General Data Protection Regulation - **GDPR**)<sup>8</sup> entered into force on the 25<sup>th</sup> of May 2018. The GDPR, whose legal basis has been identified in Art 16 (2) of the Lisbon Treaty (TFEU)<sup>9</sup> regulates the protection of individuals about the processing of personal data and the free movement of such data.

The GDPR replaces the Directive 95/46 on the protection of individuals about the processing of personal data and the free movement of such data.<sup>10</sup> The EU policymakers deemed to move from a directive, which has to be implemented by the Member States through national legislation, to a regulation – the GDPR –which is directly applicable across all EU Member States. The main reason for this shift was that the Directive 95/46/CE *‘ha[d] not prevented fragmentation in the implementation of data protection across the Union’*. Also, achieving *‘effective protection of personal data requires the strengthening and the setting out in detail of the rights of data subjects and the obligations of those who process and determine the processing of personal data’*.<sup>11</sup>

The GDPR represents the cornerstone of the new Data Protection Framework, it sets a higher standard with respect to the protection of individuals. It is expected that the new regime will enhance business opportunities for EU entities and, consequently, boost the EU Digital Single Market. By broadening the scope of the Directive 95/46, the GDPR allocates duties and responsibilities, provides indications to data subjects on how to exercise their rights, clarifies which national law applies and consequently the national supervisory authority that should be competent in a case involving the processing of personal data.

Even if the GDPR is directly applicable, Member States keep a limited margin of appreciation in the implementation of determined matters (such as age threshold governing child’s consent, the processing of

---

<sup>7</sup> During the TRUSTS trials, processing of personalized data will occur. This would be enabled by the processing, where appropriate e.g. in UC2, of aggregated and anonymous data. Please note that data protection law does not apply to anonymous data. On the flip side, when further processed, anonymous data may enable re-identification of individuals, in which case data protection law would apply. For this reason, anonymization services are provided to prevent re-identification.

<sup>8</sup> Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] *O.J.E.U.*, L119/1.

<sup>9</sup> Art 16(2) TFEU: *The European Parliament and the Council, acting in accordance with the ordinary legislative procedure, shall lay down the rules relating to the protection of individuals with regard to the processing of personal data by Union institutions, bodies, offices and agencies, and by the Member States when carrying out activities which fall within the scope of Union law, and the rules relating to the free movement of such data. Compliance with these rules shall be subject to the control of independent authorities.*

<sup>10</sup> Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and the free movement of such data [1995] *O.J.E.U.*, L281/31.

<sup>11</sup> Rec 9 -11 GDPR.

sensitive data, the exceptions to some of the data subject's rights, the provisions dealing with the data protection officer and the rules on data transfers).<sup>12</sup>

In the context of the TRUSTS project, the initial question stemming from the privacy and data protection framework, is whether the data collected and/or processed by relevant stakeholders in the project, can be qualified as personal data, and if they do, which rules should be applied.

GDPR requirements:

#### LR1 GDPR compliance is required for personal data

##### Remarks

The list of data protection principles includes:

- Lawfulness, fairness, and transparency.
- Purpose limitation.
- Data minimization.
- Accuracy.
- Storage limitation.
- Integrity and confidentiality.
- Accountability.

Data controllers, while performing their activities, have to comply with these principles.

## 5.2 e-Privacy Regulation Proposal

The EC, aware of the rapid development and an ever-growing volume of data used in the electronic communications environment, has decided to amend the e-Privacy directive and is currently working on a new e-Privacy Regulation (EPR).<sup>13</sup> The EC has focused its action on broadening the scope of the Directive regulating entities that were not regulated by the ePD. The proposed text aims at reinforcing the regime of protection for users and subscribers of electronic communications services, as part of the EU Digital Market Strategy. In particular, the text aims at ensuring proper enforcement by introducing new compliance obligations and sanctions in situations of non-compliance.<sup>14</sup> Also, the proposed Regulation intends to update the current global standards regarding the confidentiality of communications.<sup>15</sup> Taking into account the development in this field, the new EPR should be considered complementary to the GDPR. Indeed, contrary to the ePD, and similarly to the GDPR, the proposed EPR will be directly applicable within the EU.

<sup>12</sup> See Winfried Veil's map on the opening clauses in the GDPR <<https://www.flickr.com/photos/winfried-veil/24134840885/in/dateposted>> accessed 10 May 2019.

<sup>13</sup> Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Safeguarding Privacy in a Connected World A European Data Protection Framework for the 21st Century, COM(2012) 9 final, 25 January 2012, available at [http://ec.europa.eu/justice/data-protection/document/review2012/com\\_2012\\_9\\_en.pdf](http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_9_en.pdf)., accessed 24 April 2019.

<sup>14</sup> Factsheet on Data Protection Reform, Why do we need an EU data protection reform?, available at [http://ec.europa.eu/justice/data-protection/document/review2012/factsheets/1\\_en.pdf](http://ec.europa.eu/justice/data-protection/document/review2012/factsheets/1_en.pdf)., accessed 24 April 2019.

<sup>15</sup> Hunton & Williams, EU Data Protection Regulation Tracker, available at <https://www.huntonregulationtracker.com/>, accessed 24 April 2019.

The GDPR is a horizontal legislation that seeks to ensure the protection of data subjects' fundamental rights regardless of the specific sector where the processing of personal data occurs. The EPR is a vertical legislation that complements and particularises GDPR's rules in the context of electronic communications services. Contrary to the GDPR, the EPR covers data that are not necessarily personal. When electronic communication data are also personal data, due to its nature of *lex specialis*, EPR provisions shall take precedence.

e-Privacy requirements:

## LR2 e-Privacy Regulation compliance is required for electronic communications

### Remarks

The classification of data gathered and generated by the TRUSTS Consortium falls under the following categories:

1. Electronic communication data that fall under Art 4(3)(a) EPR but are not personal data (Art 4(1) GDPR). In this case, EPR applies to ensure the confidentiality of electronic communication data that otherwise would not be covered.
2. Personal data that fall under the definition of Art 4(1) GDPR but cannot be qualified as electronic communication data. When this situation occurs, the GDPR provisions apply.
3. Electronic communications data that fall into the definition of personal data. When this situation occurs, the EPR will prevail over the GDPR. This implies that the legal basis for the processing of data has to be found in Art 6 EPR (and not in Art 6 GDPR).

## 5.3 Free Flow of Non-Personal Data Regulation

In line with the Digital Single Market strategy, the EC published a legislative proposal on the free flow of non-personal data in 2017.<sup>16</sup> In its General Approach on the Free Flow of Non-Personal Data Regulation (FFNPR), the Council defines the EC proposal as a '*balanced compromise that gives Member States flexibility to address core public responsibilities while respecting the principles of the free flow of data.*'<sup>17</sup> The European Parliament on its side also welcomed the initiative. The Committee for the Internal Market and Consumer Protection has defined the free flow of non-personal data as the fifth freedom of the EU Single Market after goods, people, services and capitals.<sup>18</sup> After a negotiation phase between the European Parliament and the Council (under EC supervision), an overall agreement was reached, and final approval occurred at the beginning of November

<sup>16</sup> European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Region – Commission Work Programme 2016 – No time for business as usual', [https://ec.europa.eu/info/sites/info/files/cwp\\_2016\\_en\\_0.pdf](https://ec.europa.eu/info/sites/info/files/cwp_2016_en_0.pdf) accessed 24 April 2019.

<sup>17</sup> For the version proposal as revised by the Council, see: <http://data.consilium.europa.eu/doc /document/ST-15724-2017-REV-1/en/pdf> accessed 24 April 2019.

<sup>18</sup> "This regulation *de facto* establishes data as the fifth freedom on the EU Single Market. By removing borders, burdens and barriers such as data localisation rules, we enable a level playing field for European companies to compete globally. This legislation is truly a game changer, potentially providing enormous efficiency gains for both companies and public authorities. It will reduce data protectionism, which is threatening the digital economy, and pave the way for artificial intelligence, cloud computing and big data analysis".

2018. The FFNPDR was signed on the 14<sup>th</sup> November 2018, entered into force at the end of December 2018 and is applicable from May 2019.<sup>19</sup>

The EC considers the free flow of non-personal data a fundamental building-block of the Digital Single Market Strategy. According to the EC, the FFNPDR, removing the national restrictions to the free flow of non-personal data, will contribute to boosting the EU economy, generating growth of up to 4% GDP by 2020.<sup>20</sup>

The Commission has recognised four barriers to data mobility within the EU market:

- (1) Data localisation restrictions by member states' public authorities;
- (2) Obstacle put in place by IT systems' vendors;
- (3) Complex EU legal patchwork that leads to legal uncertainty;
- (4) Lack of trust due to security risks and concerns about the cross-border availability of data for regulatory purposes.<sup>21</sup>

The removal of the legal obstacles is considered preliminary not only for enhancing the economy but also for boosting innovation (with expected progress in the field of AI, IoT and autonomous systems).

Free flow of non-personal data requirements:

**LR3 Free Flow of Non-Personal Data Regulation compliance is required for data exchange between the node and stakeholders of the TRUSTS data marketplace.**

#### **Remarks**

To boost the Digital Single Market, the FFNPDR aims to remove all the barriers that are hampering the free movement of non-personal data. Doing so, the FFNPDR identifies three main actions to achieve its purpose: prohibition of mandatory data localisation requirements, guarantee data availability for competent authorities, and facilitation of data porting by users.

## **5.4 Platform-to-Business**

On 26<sup>th</sup> April 2018, the EC published its Proposal for a Regulation on promoting fairness and transparency for business users of online intermediation services (Platform-to-Business Regulation - P2BR). The Proposal was anticipated by a public consultation and a communication in 2016. On the 14<sup>th</sup> of February, the Council and the Parliament reached an overall agreement that remains to be approved. Once approved, it was published in the EU Official Journal to enter into force at the beginning of November 2019.

<sup>19</sup> Regulation (EU) 2018/1807 on a framework for the free flow of non-personal data in the European Union *O J L 303*, 28.11.2018, p. 59–68.

<sup>20</sup> Deloitte, “Measuring the Economic Impact of Cloud Computing in Europe”, final report prepared for the European Commission <https://ec.europa.eu/digital-single-market/en/news/measuring-economic-impact-cloud-computing-europe> accessed 24 April 2019.

<sup>21</sup> See, for a visualisation and information on the objectives of the proposal: European Commission, ‘State of the Union 2017 – Free flow of non-personal data’, <https://ec.europa.eu/digital-single-market/en/news/free-flow-non-personal-data> accessed 24 April 2019.



With this initiative, the EC intends to legislate in the area of business platforms, which has, so far, not been addressed by specific legislative initiatives. Considering that the final approved text is not yet available, this deliverable focuses on the proposal published by the Commission in 2018. If any substantial changes occur, they will be reported and described in the upcoming deliverables.

The P2BR is part of the legislative measures promoted by the EC for the Digital Single Market strategy. The proposal is the first legislative initiative in the field of platforms and focuses only on a specific type of platforms, namely, those offering services or products to the same users of their business clients. The P2BR foresees for them a list of measures ensuring transparency and fairness. Doing so, the EC aims to temper the natural asymmetries that characterise the relationship between platforms and their suppliers, establishing a fair and trustworthy innovation-driven ecosystem.

Platform-to-business requirements:

#### LR4 Platform-to-Business Regulation compliance is required to safeguard data marketplace operational transparency and fairness

##### **Remarks**

The P2BR follows two main principles, namely, transparency and fairness.

The P2BR foresees transparency obligation for providers of intermediation services to inform, through clear, unambiguous and readily available contractual terms and conditions, about the treatment, the criteria used to rank their products and the requirements to suspend or terminate their services.

Moreover, the P2BR aims to achieve fairness through the settlement of effective out-of-court redress mechanisms such as internal handling systems for business users and mediation procedures. To facilitate the process, contractual terms and conditions prepared by the intermediaries have to include a list of independent mediators that can be approached to settle disputes.

## 6 Questionnaires and Interviews

The aim of this deliverable is to analyse the requirements retrieved from financial institutions, telecom operators and corporate data providers/users, as well as from industrial stakeholders. The requirements' capture involves the documentation of industrial and regulatory needs and opinions about new innovative data marketplace service verticals, which will set the baseline for conducting the actual measurements during the use case trials. Requirements collection has been performed via electronic surveys and through dedicated workshops/interviews with the foreseen actors of every case study.

The survey is included in Annex I and is structured as follows:

- Inform & Consent form for GDPR compliance
- Demographics to acquire information on the organisation that the questionnaire responder represents
- Role of the interviewee within the organisation, level of related experience, authority
- Questions depending on the role e.g. data buyer, data seller, application/service provider, application/service buyer

- Questions for the preferred applications that are provided through a data marketplace and the related processes
- Questions about the regulatory status with respect to data marketplaces and potential gaps
- Questions with respect to preferred pricing schemes that a data marketplace should offer
- Question for the willingness of the interviewee to be interviewed

In order to maximise participation all the TRUSTS partners the questionnaire has been forwarded not only to the partners that participate in the task 2.2 but also to relevant employees within their organization and external collaborators. In the sequence, each partner performed face2face interviews.

The guidelines for the interviews are presented in Annex II.

Due to the limited amount of time to perform a statistically correct pan-European survey we focused on targeting selective relevant interviewees able to provide key insights for the required data marketplace usage and offered services<sup>22</sup>.

In the following section there exists an in-depth analysis of the questionnaire and interview responses.

## **7 Stakeholders feedback analysis for a commercial financial and operators' industry vertical data marketplace platform**

In this section an in-depth quantitative and qualitative analysis of the feedback to the questionnaire is achieved.

### **7.1 Electronic Survey**

#### **7.1.1 Method and Procedure**

Following the preparation of the questionnaire, the corresponding electronic survey was made available through the SurveyGizmo<sup>23</sup> platform, which features a GDPR-compliant data centre and has in place tools for data privacy disclosures and opt-in statements in surveys. The survey link<sup>24</sup> was disseminated to all TRUSTS partners, who were asked to further disseminate it to an as-wide-as-possible audience. The survey remained online for a total of 60 days. After that, responses were filtered to eliminate invalid ones, before proceeding with their analysis.

---

<sup>22</sup> Also COVID-19 lockdown had a negative impact on the process. Nevertheless, the consortium managed to schedule remote sessions with key interviewees.

<sup>23</sup> <https://www.surveygizmo.com/>

<sup>24</sup> <https://www.surveygizmo.eu/s3/90212977/TRUSTS-requirements-elicitation-questionnaire>

### 7.1.2 Participants

A total of 31 individuals responded to the electronic survey, however as questions were not mandatory, the total responses per question are sometimes less than the number of respondents.

The majority of participants (21) originate from the private sector, while a considerable percentage (6) works in Academic organizations and a smaller one (4) in the public sector (**Fehler! Verweisquelle konnte nicht gefunden werden.** left). Input provided by participants from the private sector, identified that their organizations fall into the following domains: Telecommunications, Information Technology, Audit and Accounting, Financial Services, and Legal Services. In addition, participants' organizations vary in size, including nearly equal numbers of small (1-50 employees), medium (51-20 employees), medium-large (251-1000 employees) and very large (more than 1000 employees) organizations (Figure 2 right).

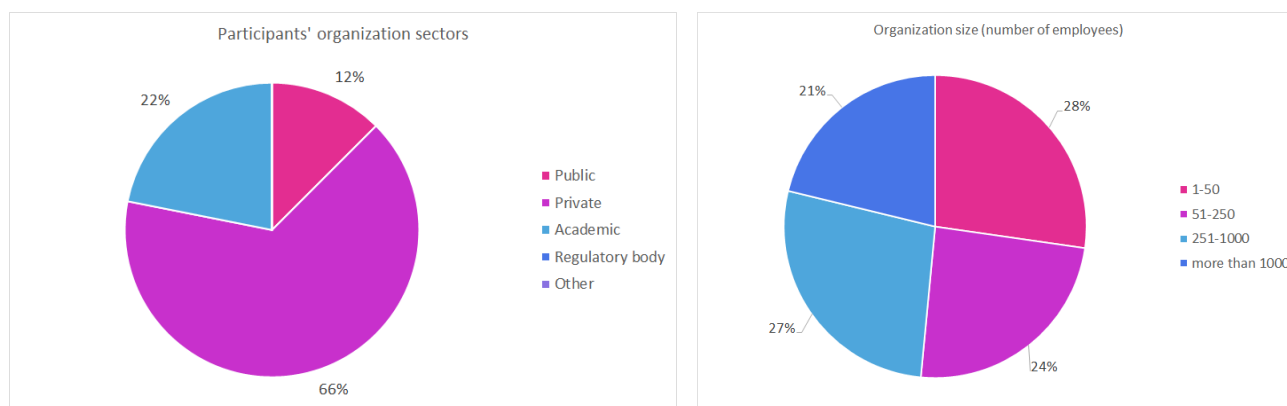


Figure 2: Participating organizations' information

Regarding the roles of the participants' organizations in the data exchange value chain (Figure 3), participants were asked to select multiple options if applicable, as an organization might be for example involved both as a data buyer and as a data provider. Out of the total of 55 responses<sup>25</sup>, it turns out that a considerable proportion of organizations is involved in the data exchange value chain as application or service providers (45.16%), followed by equally high proportions of data buyers (38.71%) and providers (35.48%). Smaller percentages were recorded for application or service users (16.13%), standardization or regulation bodies (9.68%) and data marketplace platform operators (3.23%, only one organization). A considerably high number of responses was also recorded for the other option, in which clarifying free-text answers of the participants revealed the following key roles: research in data science and data exchange value chain, market data evaluation, as well as data privacy and protection.

<sup>25</sup> Each respondent could declare more than one roles (i.e, data provider and data buyer), thus the total number of responses is greater than the number of respondents. This is also holds for all the below mentioned statistical analysis.

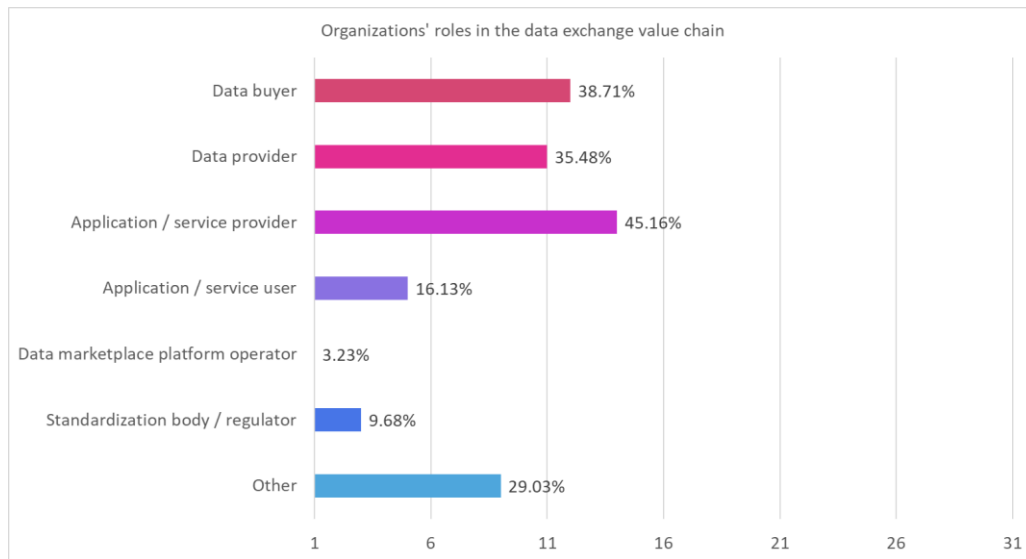


Figure 3: Roles of participants' organizations in the data exchange value chain

Regarding participants, nearly half (45.16%) characterized themselves as domain experts, followed by considerable proportions of business drivers (38.71%) and technical drivers (35.48%), and a smaller percentage of strategical drivers (19.35%), as illustrated in (Figure 4 left). Regarding their level of management, a considerable proportion of participants are top-level of management employees (administrative officer 35.48%), while smaller but nearly equal percentages middle-level (executive officer 19.35%), lower level (operative officer 16.13%), and researchers (19.35%) (Figure 4 right). A smaller percentage selected the "Other" option (9.68%), with only one out of these three individuals further clarifying their role as developers. Furthermore, the majority of participants have many years of business experience, with 71% in total working in the field for more than 10 years. In detail, 35.48% have more than 20 years of experience, 35.48% 10 to 20 years, 16.13% 5 to 10 years, 9.68% 2 to 5 years and only 3.23% (one participant) less than 2 years (Figure 5).

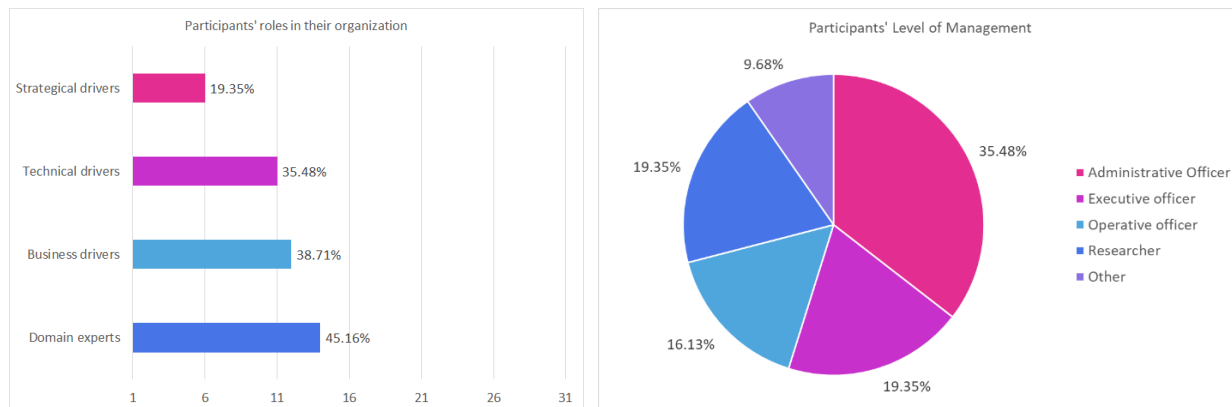


Figure 4: Participants' roles and level of management in their organizations

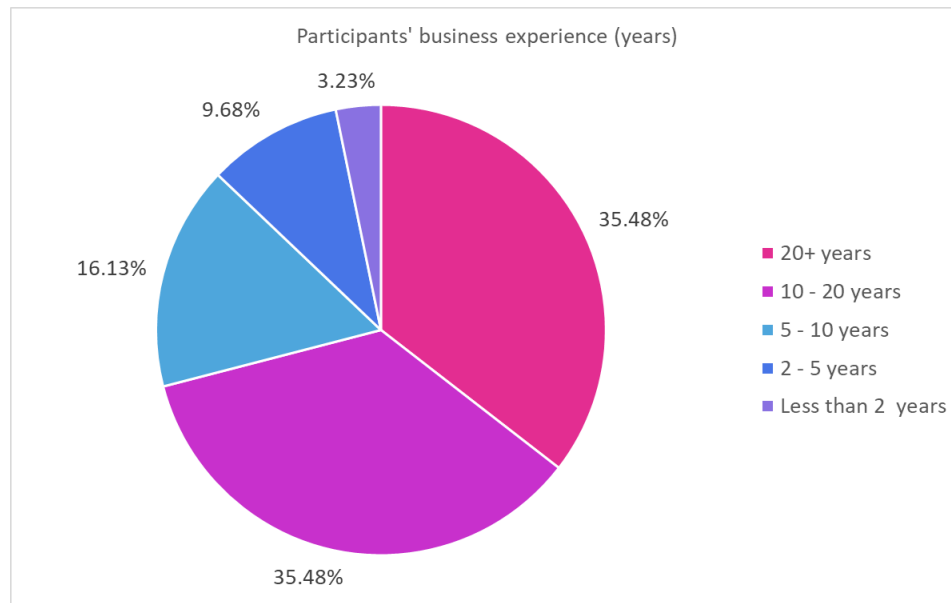


Figure 5: Participants' business experience

Regarding their relationship with the data sales and / or buying processes of their organization (Figure 6), it is noteworthy that although nearly half of the participants declared none or low involvement, only 10% stated none or low understanding of the processes. In particular, regarding their understanding of the relevant processes, as self-assessed on a scale from 1 (no understanding at all) to 5 (full understanding), 6.45% declared no understanding (rating 1), 3.23% low (rating 2), 32.25% moderate (rating 3), 29.03% high (rating 4), and 22.58% full understanding (rating 5). In terms of their involvement, 35.48% indicated now involvement (rating 1), 16.12% low (rating 2), 6.45% moderate (rating 3), 22.58% high (rating 4) and 12.90% full involvement (rating 5).

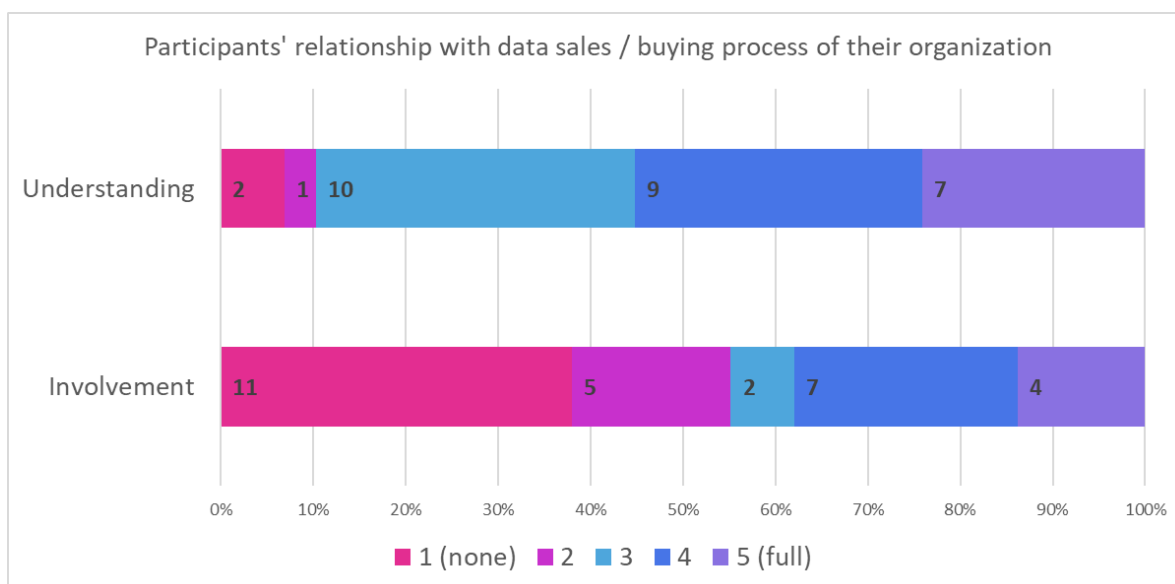


Figure 6: Participants understanding and involvement in data sales / buying process in their organization

### 7.1.3 Results

Participants, who indicated that their organization is a data buyer, were asked to further elaborate on three questions: (i) how often they buy data (ii) what type of data they buy and (iii) what is the desired process to buy data. The frequency of buying and type of data bought is clearly dependent on the needs of each organization (Figure 7). In particular, half of the data buyers indicated that their organizations buy data once or twice per month. The remaining responses were as follows: 16.67% (2 organizations) buy data quarterly, 16.67% rarely, 8.33% (1 organization) five times per month, and another 8.33% on a need basis, according to their day-to-day operations. The types of data bought are market research (33.33% of the buyers), corporate (33.33%), financial (25%), geospatial (16.67%), benchmarking (16.67%), AI/ML (8.33%), CRM (8.33%), and other open data (25%) such as public listing, meteorological, and marine traffic.

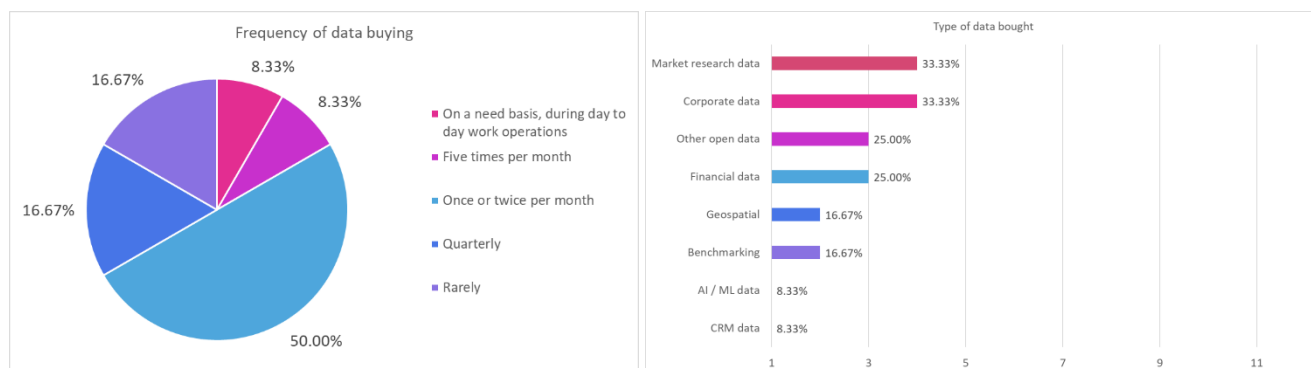


Figure 7: Data buying frequency and type

Regarding the desired process to buy data, participants highlighted a number of qualities and user requirements that they would like to be accomplished. In particular, the following requirements were identified:

- The process should be accomplished online, through a one-stop-shop service.
- The process should be confidential.
- The process should be standardized.
- The process should provide easy access.
- Data discovery should be based on the FAIR principle, being Findable, Accessible, Interoperable, and Reusable.
- The process should be accessible upon login on a centralised look up service.
- Searching for datasets should be easy.
- Searching datasets through keyword should be supported.
- Before paying, information on the dataset cost should be provided.
- Additional information about the dataset that should be provided: data description, and owner.
- The data should have straightforward mapping with the required business entities.
- The data should be fully anonymized.
- Payment should be easy and fast (one-click process).
- After the payment, direct dataset downloading should be supported.
- Various subscription schemes should be supported, such as annual license subscription.
- It should be possible to meet with data providers and choose the best solution / partnership.
- The platform should support online contracts.

Participants, who indicated that their organization is a data seller, were asked to further elaborate on three questions: (i) how often they provide or sell data (ii) what type of data they provide or sell and (iii) what is the desired process to provide or sell data.

Regarding data selling frequency (Figure 8 left), it is noteworthy that a considerable proportion of organizations sell data daily (44.44%) or a few times per month (33.33%). Quite often (ten times) in a month – but not daily – was also provided as response by 11.11%, while a yearly selling frequency was reported by another 11.11%.

The types of data that organizations typically sell exhibited narrow convergence among respondents (in total 10 responses were analysed as valid). In particular, 30% sell financial data, 20% CRM, 10% Geospatial, 10% corporate, 10% market data and 20% other open data (Figure 8 right).

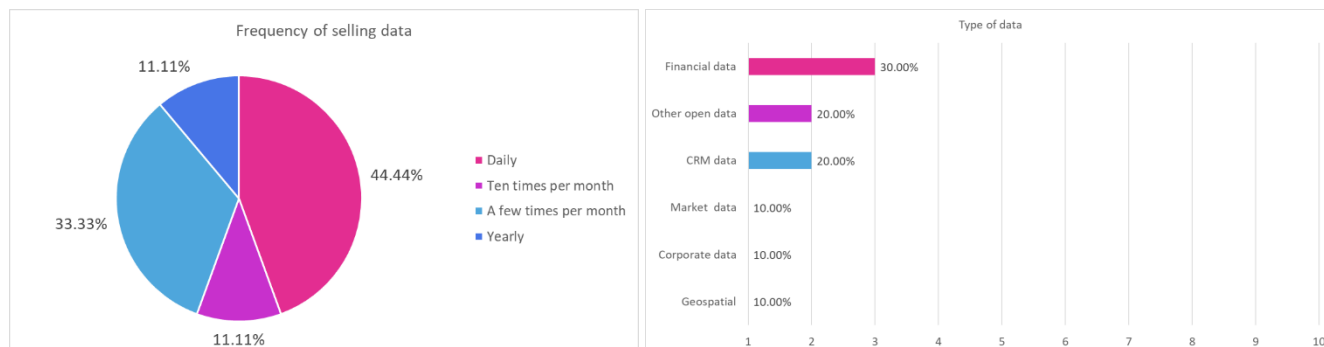


Figure 8: Data selling frequency and type

Regarding the desired process for selling data, participants highlighted a number of qualities and functional requirements that they would like to be addressed. In particular, the following requirements were identified:

- The process should preferably be unified and standardized.
- The entire process should be electronically supported.
- The process should be GDPR compliant and approved.
- The selling platform should support subscription, including annual subscription schemes.
- It should be easy to describe the provided datasets, through on line metadata forms following standards, e.g., Dublin Core.
- One-click data uploading and confirmation of successful accomplishment (like file transfer services).
- The platform should provide intuitive insights with regard to the provided data (e.g., number of downloads per different time periods, geographical distribution, etc.), as well as online financial management of data sells.
- For providing datasets to specific application / service providers, the process should involve selecting the specific provider, announcing the data, and – upon agreement - contracting the provider.
- The platform should support direct dataset requests from clients.
- The platform should support filtering of datasets and a corresponding classification, to free and paid.
- The platform should support open data.

Participants who indicated that their organization is a data application or service provider, were asked to further elaborate on three questions: (i) how often they provide data applications or services (ii) what type of applications or services they provide (iii) what is the desired process to provide applications or services.

Regarding the frequency of providing services (Figure 9 left), respondents indicated that their organizations either provide such services on a daily basis (66.67%), or a few times per month (33.33%). In addition, the types of provided services were identified as mainly financial (38.46%), data analytics (38.46%) and Information Technology or Software solutions (38.46%) (Figure 9 right). Additional services include telecommunications (7.69%), consulting and training (7.69%), as well as market analysis (7.69%).

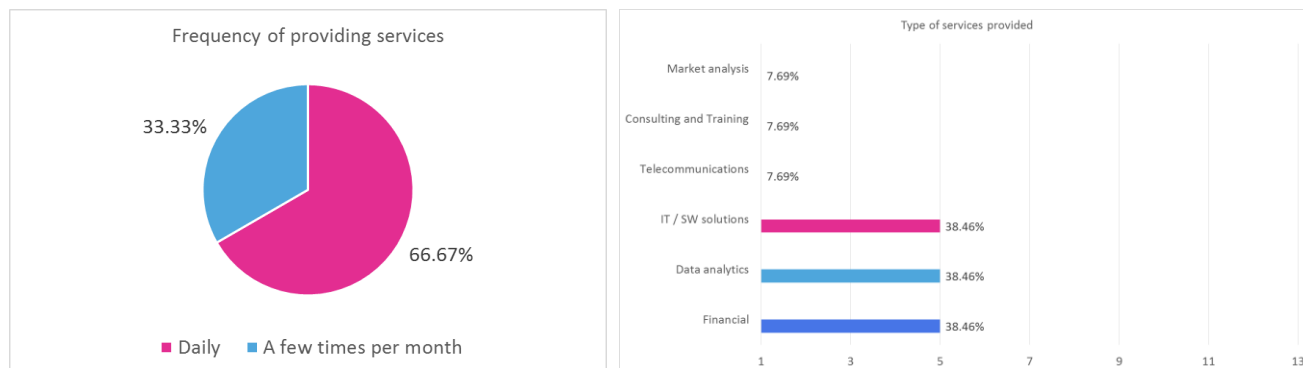


Figure 9: Frequency and type of provided data applications or services

Regarding the desired process for providing services, participants highlighted the following requirements:

- The process should be electronic.
- The process should be confidential, according to GDPR policies.
- Open data should be supported.
- Providing services directly to end-customers should be supported.
- The platform should support subscription, featuring annual license subscription as well.
- A connection of the platform with highly visiting applications marketplaces, such as Google Play and Apple Store, should be provided<sup>26</sup>.
- Retrieving datasets should be easy.
- Keyword based searching of datasets should be supported.
- Alternatively to keyword searching for a dataset, browsing through structured content categories should be supported.
- Each dataset should include description and tags.
- Ratings and comments from other users who have already used the dataset should also be provided.
- Information about the anonymization of the dataset is important.
- Viewing a small sample of the dataset before buying it would also be useful.
- A discrete distinction between free and paid datasets should be provided.
- Networking between partners should be supported.

Participants, who indicated that their organization buys data applications or services, were also asked the similar questions, namely (i) how often they buy such applications or services, (ii) what applications or services do they buy and (iii) what is the desired process for doing so.

Regarding the frequency of such transactions (Figure 10 left), respondents answered that it is annual (33.33%), monthly (a few times per month, 50%), or regular (almost daily, 16.67%). The types of applications or services bought (Figure 10 right) are mostly marketing (50%), financial (33.33%), and human resources (16.67%).

<sup>26</sup> All respective decisions within TRUSTS will take into account GDPR compliance.



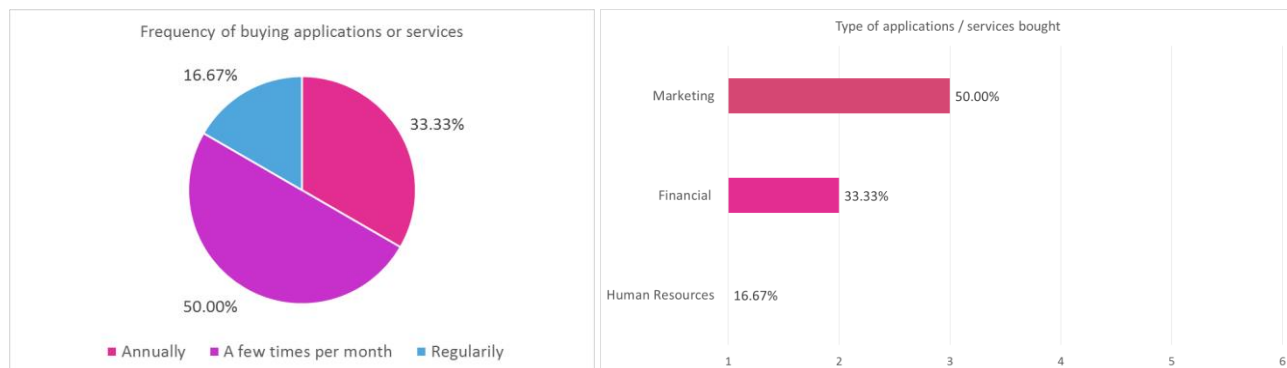


Figure 10: Frequency and type of data applications or services bought

Regarding their requirements for the process of buying applications or services, respondents indicate that they would like to:

- Preview the service before buying it.
- Search via keyword, but also by providing other options such as the cost they have to pay.
- Review the service specifications.
- Carry out comparisons between services from different providers.
- Review feedback on the user experience from other customers.
- Be able to contact vendors directly.

Next, participants who indicated that their organization is a data marketplace platform operator were also asked (i) how often they provide services through their platform (ii) what applications or services do they provide and (iii) what is the desired process for doing so.

Only two participants responded to this set of questions, therefore their answers are not analysed through diagrams. Regarding the frequency, participants identified that they provide such services daily. The type of services provided pertained to open data, as well as online sales and marketing. Regarding the desired process for providing services, no specific requirements were identified, besides support for open data and advanced search mechanisms.

The next set of questions pertained to participants who indicated that their organization is a standardization body / regulator. First, they were asked to describe the standardization status with respect to data exchange between different organizations and data marketplaces. Responses in the question highlighted the following:

- Standardization is currently limited.
- IDSA co-developed (with its members) the IDS Reference Architecture 3.0 IDS (RAM3.0) (IDS Ram4.0 is to be released during Hannover Messe 2020). The IDS RAM is providing architecture specifications defining the individual components of the International Data Space (Connector, Broker, App Store, etc.) in detail.
- IDSA initiated the release of the German standard (DIN Spec 27070) for exchanging industrial data. DIN SPEC 27070 specifies the requirements to be met by a security gateway used for exchanging industrial manufacturing data and services, including the environment such a gateway can be used within. An international standard is in progress.

Following, participants were asked to identify in their opinion the standardization gaps and the way forward to boost the data marketplace endeavour, and also to describe the required standardization for federated data marketplaces. The gaps and problems identified were as follows:

- There are currently too many marketplaces and no overview.

- A standard meta-model for data exchange is missing, containing for instance standard vocabulary (e.g. Asset Administration Shell).
- Usage Control and legal framework (e.g. contracts) for data exchange is missing.

The requirements that were identified in this respect included:

- Strong authentication mechanisms to create trust.
- Intelligent matchmaking mechanisms, facilitating users to identify the data or services needed.
- Advanced searching options, including filters for the cost of a dataset or application / service.

The next questionnaire section addressed all participants, regardless of the type of their organization and pertained to data marketplaces.

First, participants were asked to select from a list of options the functions that they would like to be provided by a data marketplace platform. A total of 23 participants responded to this question, revealing the most important features of a data marketplace (Figure 11). There was no single answer that remained unselected by at least five participants; therefore they all constitute requirements of the platform. In particular, the features order by preference are as follows: anonymization (73.91%), user role rights according to GDPR processes (69.75%), datasets catalogue (69.75%), user authentication (65.22%), billing (56.52%), standardization information (52.17%), datasets discovery services (52.17%), transaction logs (47.83%), metadata hosting (47.83%), datasets / applications trading service (43.48%), datasets rating (43.48%), datasets valuation (39.13%), third party applications catalogue (34.78%), de-anonymization risk analysis (30.43%), multiparty computation (26.09%), private set intersection (21.74%), and data hosting (21.74%).

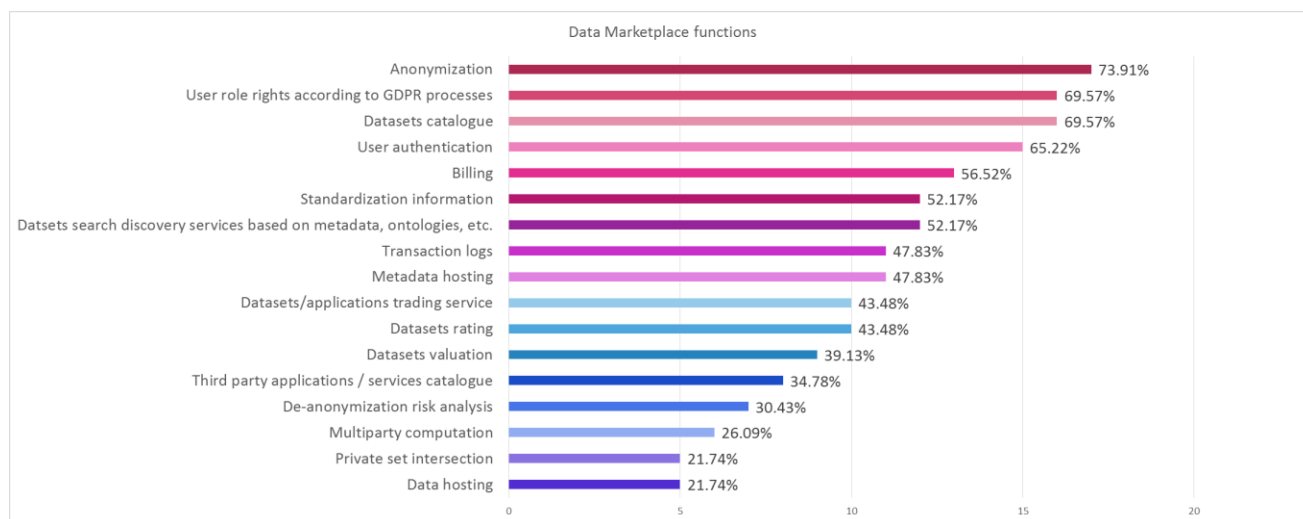


Figure 11: Data marketplace functions

A chi square test of independence showed that there was no significant association between the type of organization and the preference over specific data marketplace functions,  $\chi^2(32, 31) = 11.96$ ,  $p = .74$ . In addition, there was no significant association between the role of an individual (i.e. domain expert, executive officer, operative officer, researcher) with the data marketplace functions selected,  $\chi^2(48, 31) = 39.34$ ,  $p = .80$ . Last, a chi square test of independence indicated that the null hypothesis that the role of one's organization is independent from their preference regarding the data marketplace functions is rejected,  $\chi^2(96, 31) = 148.18$ ,  $p = .0005$ . Further analysis through chi-square tests for the effect of each variable on each marketplace function revealed that:

- For anonymization, a high percentage of application / service providers denoted that it is a desired functionality, exhibiting a statistically important difference,  $\chi^2(1, 31) = 5.8$ ,  $p = .01$  (Figure 12)

- None of the respondents who represented data buyers selected de-anonymization risk analysis as a desired function, a percentage exhibiting a statistically important difference with the responses from other organization categories,  $X^2(1, 31) = 4.69$ ,  $p = .03$  (Figure 13)
- Regarding the function of datasets catalogue, high percentages of respondents from data buyers and application service providers exhibited a preference as a desired function, which was statistically important (data buyers:  $X^2(1, 31) = 4.69$ ,  $p = .03$ , application / service providers  $X^2(1, 31) = 7.42$ ,  $p = .006$ ). A low percentage, statistically important, was observed for the other category,  $X^2(1, 31) = 4.38$ ,  $p = .03$  (Figure 14).
- For datasets valuation, half of the application / service providers group denoted that it is a desired functionality, exhibiting a statistically important difference,  $X^2(1, 31) = 5.44$ ,  $p = .02$  (Figure 15).
- Datasets search discovery was denoted as a desired function by a rather high percentage of application / service providers, exhibiting a statistically important difference,  $X^2(1, 31) = 7.03$ ,  $p = .008$  (Figure 16).
- Standardization information was deemed as a highly desired function by respondents of application / service user organizations,  $X^2(1, 31) = 4.28$ ,  $p = .03$  (Figure 17).

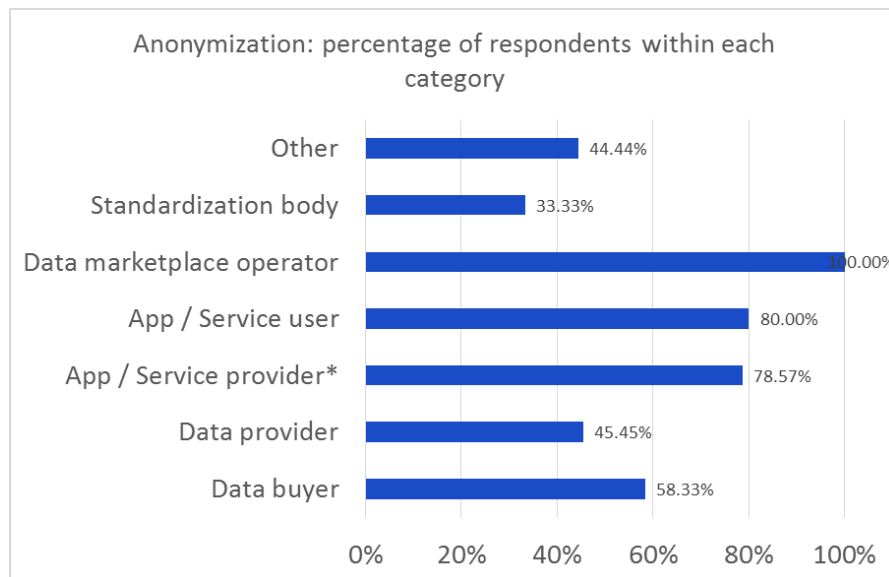


Figure 12: Anonymization (desired functionality) – percentages of responses within organization roles' categories

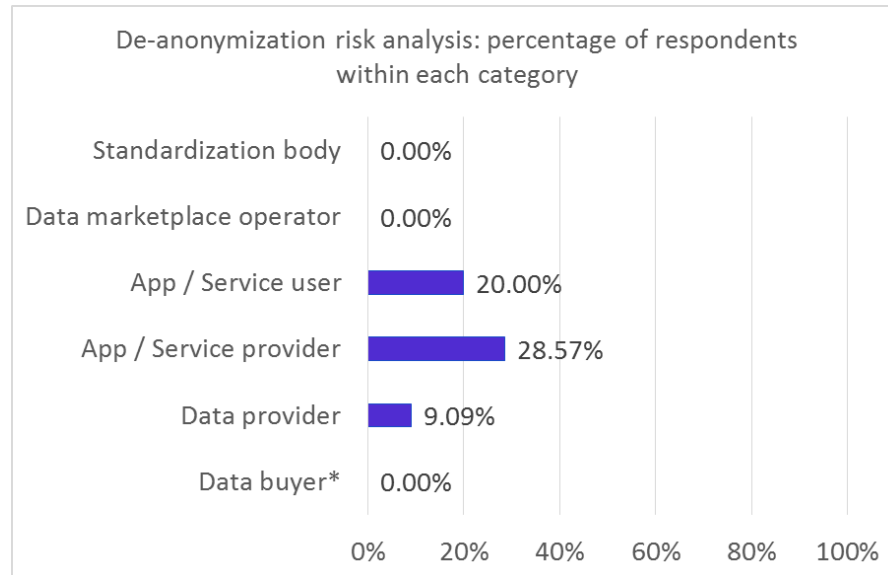


Figure 13: De-anonymization risk analysis (desired functionality) – percentages of responses within organization roles' categories

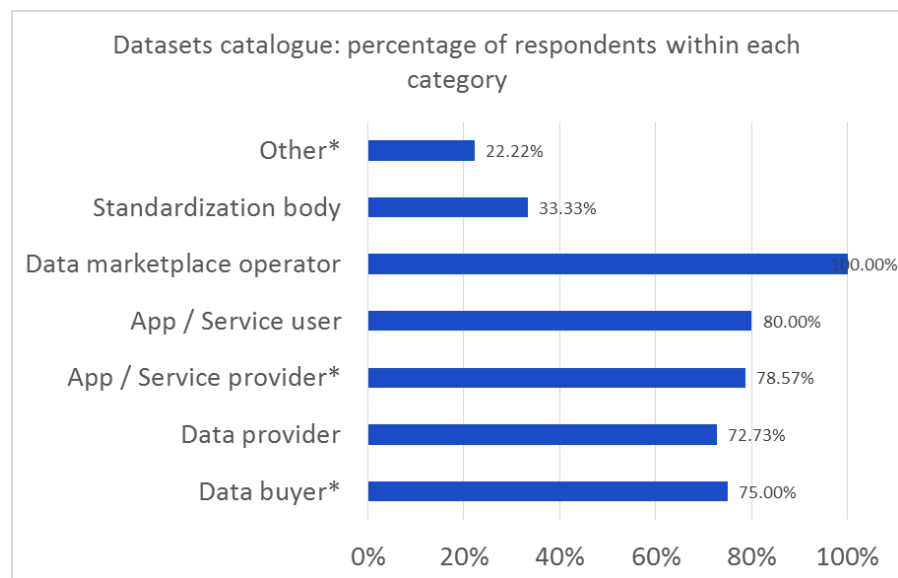


Figure 14: Datasets catalogue (desired functionality) – percentages of responses within organization roles' categories

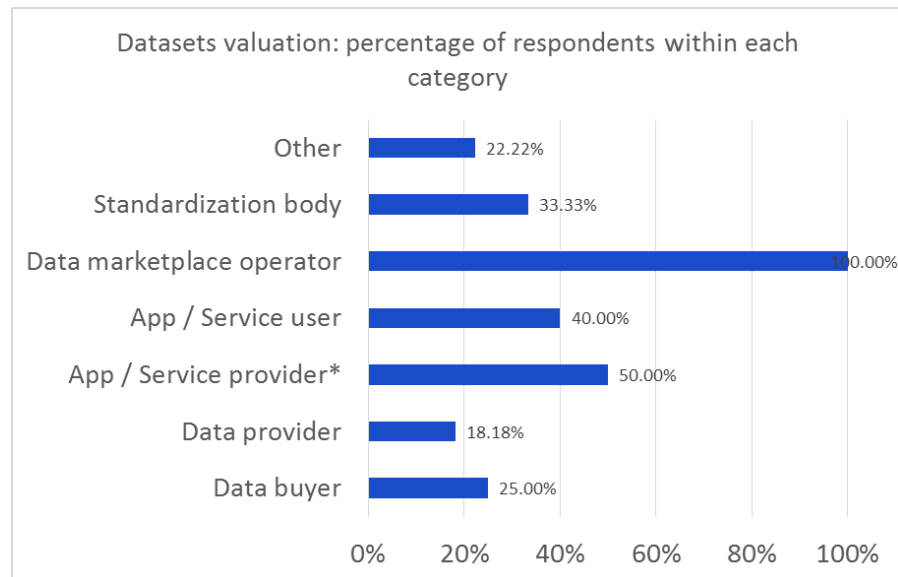


Figure 15: Datasets valuation (desired functionality) – percentages of responses within organization roles' categories

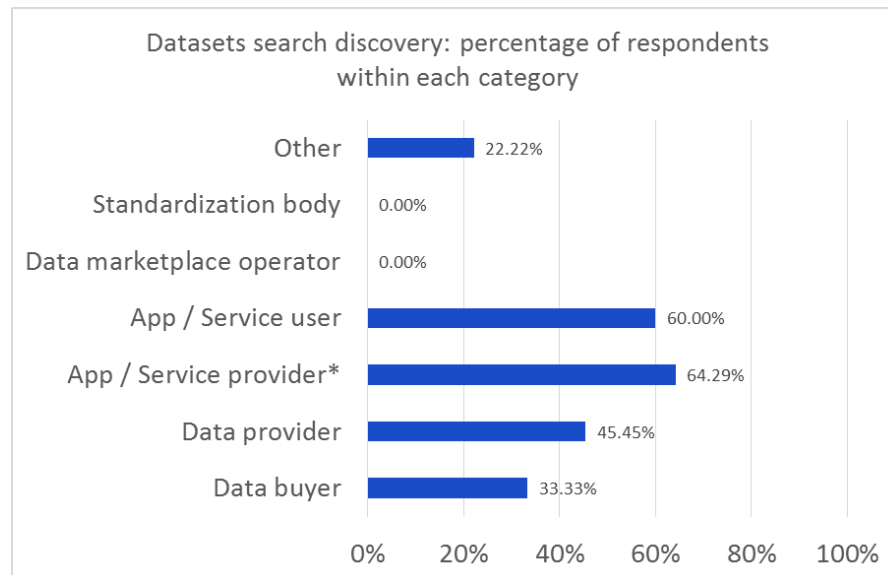


Figure 16: Datasets search discovery (desired functionality) – percentages of responses within organization roles' categories

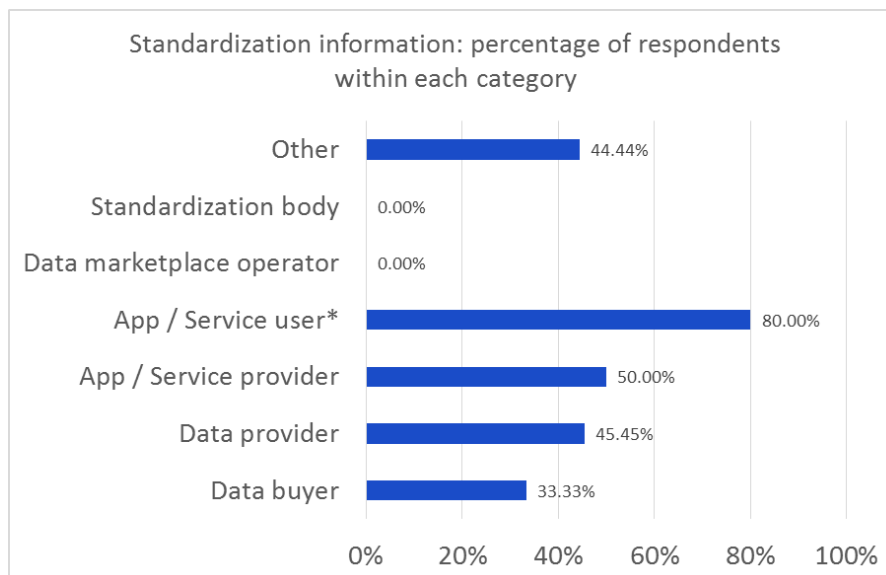


Figure 17: Standardization information (desired functionality) – percentages of responses within organization roles' categories

It is notable that respondents from standardization bodies did not report standardization information as a desired functionality. Although this finding is quite controversial, it cannot be considered as important due to the limited number of respondents from standardization bodies, which is further verified by the fact that the statistical analysis did not highlight this as a statistically important finding.

The next question asked participants to identify which functionality (from the previously identified options), they consider being mandatory. Figure 18 illustrates the comparison between desired and required for each proposed function. Overall, it can be seen that there were specific core functions that were deemed as not only desired, but also required by nearly the same amount of participants. These functions when small differences in percentages were observed (less than 10 units difference in the percentage of selected responses): anonymization, user authentication, metadata hosting, datasets valuation, de-anonymization risk analysis, private set intersection, and data hosting. Less unanimity (between 10-20 units of difference) was observed for functions that are more specialized, such as: user role rights according to GDPR processes, billing, transaction logs, trading service, datasets rating, third party applications / services catalogue, and multiparty computation. This is due to the fact that from one hand these services are quite specialized and might not be familiar to all respondents, and on the other hand that although desired they are not required for the operation of a data marketplace. Finally, considerable differences were noted for a few functions and namely, datasets catalogue, standardization information, datasets search discovery services based on metadata. The most possible reason for these differences is that these functions do not constitute prerequisites for the functionality of a data marketplace, although they are desired as a characteristic that would further enhance its overall user experience.

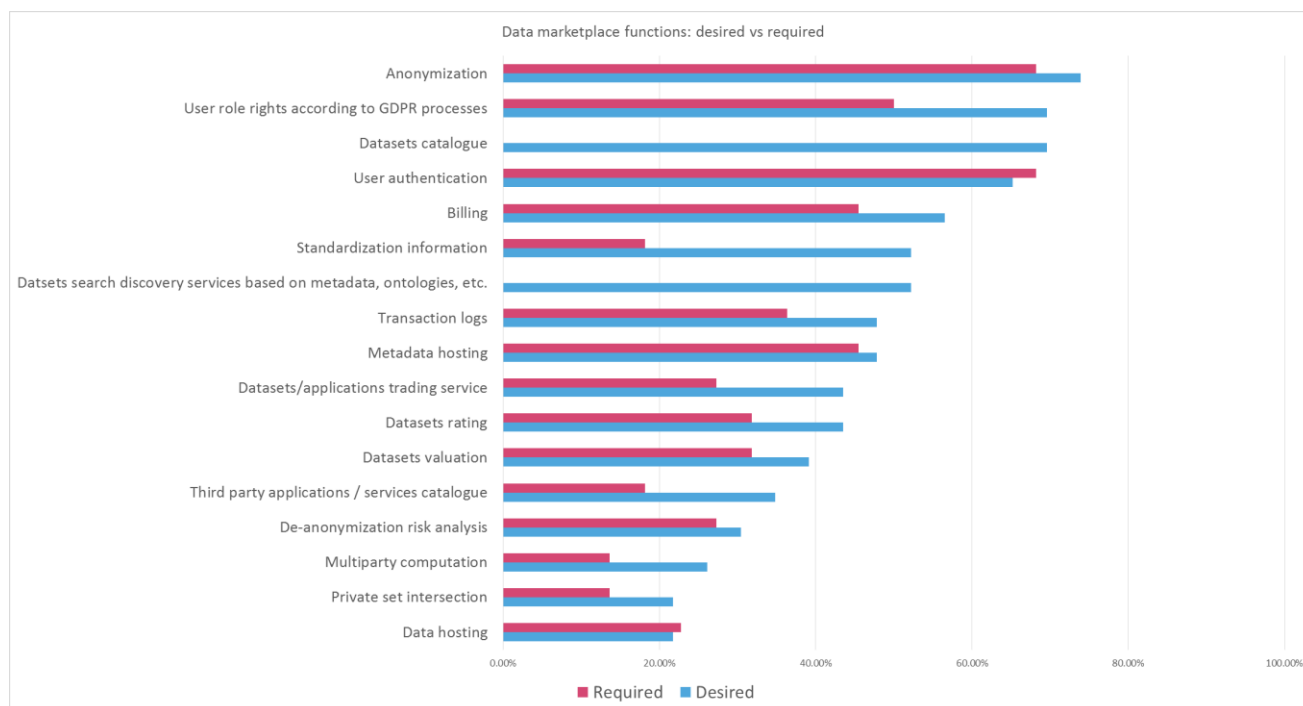


Figure 18: Data marketplace functions: desired vs. required

A chi square test of independence showed that there was no significant association between the type of organization and the preference over specific data marketplace functions,  $X^2(32, 31) = 14.37$ ,  $p = .99$ . In addition, there was no significant association between the role of an individual (i.e. domain expert, executive officer, operative officer, researcher) with the data marketplace functions selected,  $X^2(48, 31) = 37.97$ ,  $p = .84$ . Last, a chi square test of independence indicated that the null hypothesis that the role of one's organization is independent from their preference regarding the data marketplace functions is rejected,  $X^2(96, 31) = 232.48$ ,  $p < .0005$ . Further analysis through chi-square tests for the effect of each variable on each marketplace function revealed that:

- De-anonymization risk analysis was deemed as mandatory by a low percentage of application / service providers, exhibiting a statistically important difference with regard to responses from the other categories:  $X^2(1, 31) = 4.03$ ,  $p = .04$ . The exact same was observed for multiparty computation. (Figure 19, Figure 20)
- Regarding the function of user authentication, high percentages of respondents from data buyers and data providers exhibited identified it as mandatory function, a preference which was statistically important (data buyers:  $X^2(1, 31) = 5.55$ ,  $p = .01$ , application / service providers  $X^2(1, 31) = 4.04$ ,  $p = .04$ ) (Figure 21)
- The function of dataset valuation was considered by nearly half of application / service providers as mandatory at a statistically significant level:  $X^2(1, 31) = 6.00$ ,  $p = .01$ . On the contrary, none of data providers denoted it as mandatory:  $X^2(1, 31) = 4.97$ ,  $p = .02$ . The exact same percentages and preferences were noted for the function of dataset rating (Figure 22, Figure 23).
- Multiparty computation was deemed as mandatory by a low percentage of application / service providers, exhibiting a statistically important difference with response from the other categories:  $X^2(1, 31) = 4.03$ ,  $p = .04$  (Figure 24).

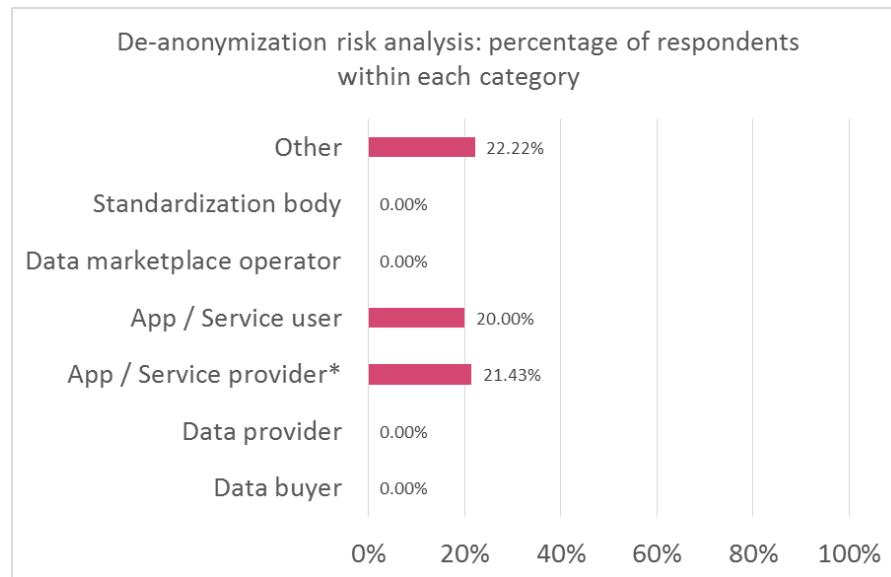


Figure 19: De-anonymization risk analysis (mandatory) – percentages of responses within organization roles' categories

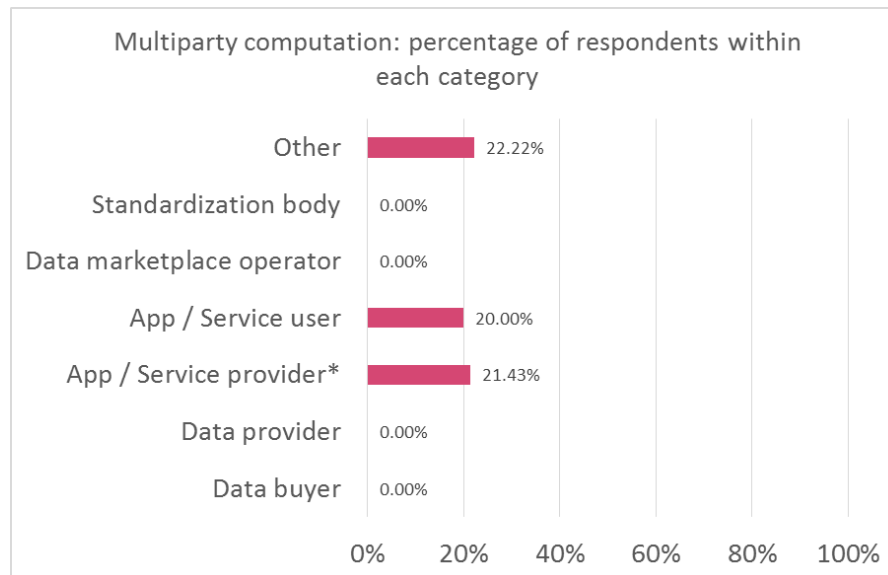


Figure 20: Multiparty computation (mandatory) – percentages of responses within organization roles' categories



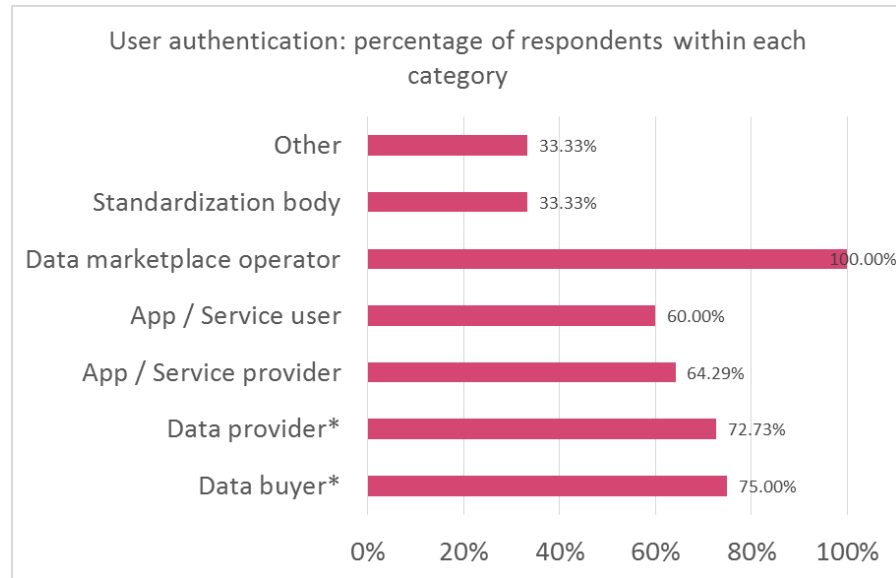


Figure 21: User authentication (mandatory) – percentages of responses within organization roles' categories

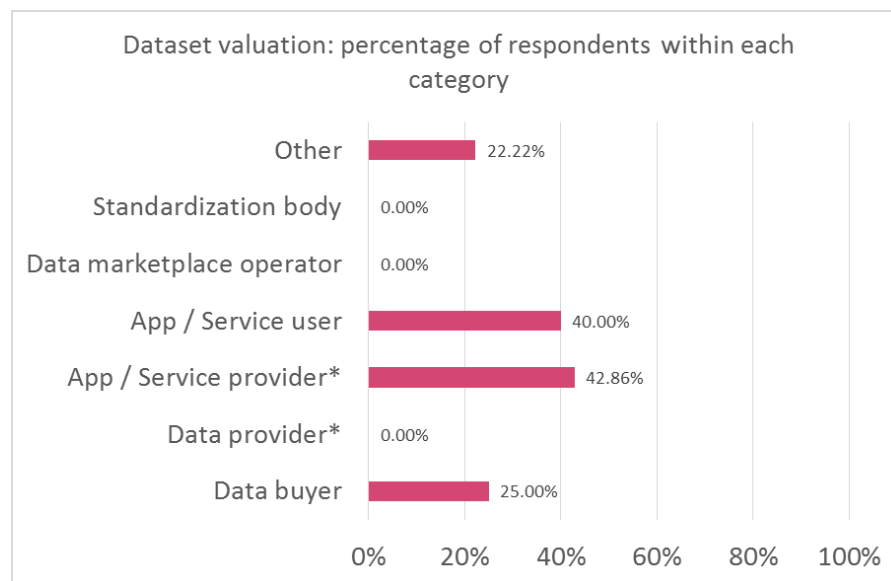


Figure 22: Dataset valuation (mandatory) – percentages of responses within organization roles' categories

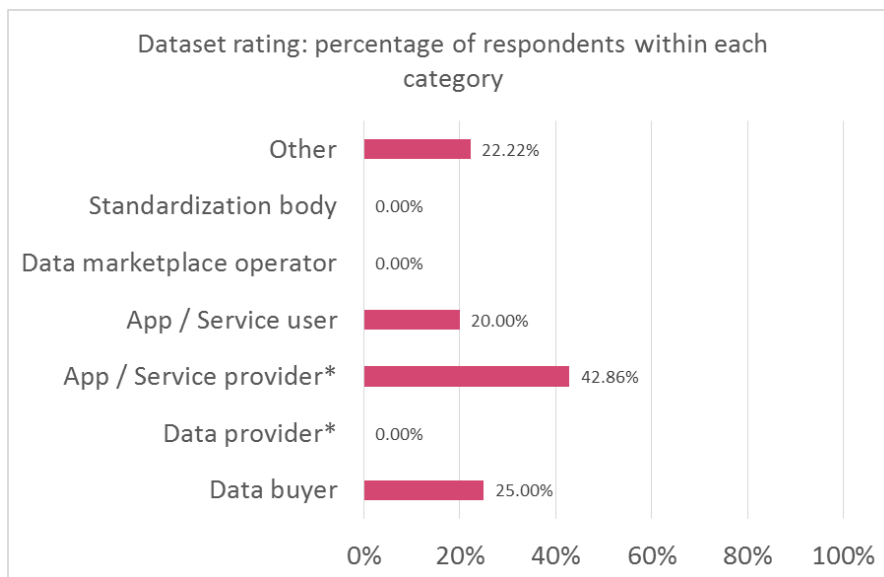


Figure 23: Dataset rating (mandatory) – percentages of responses within organization roles' categories

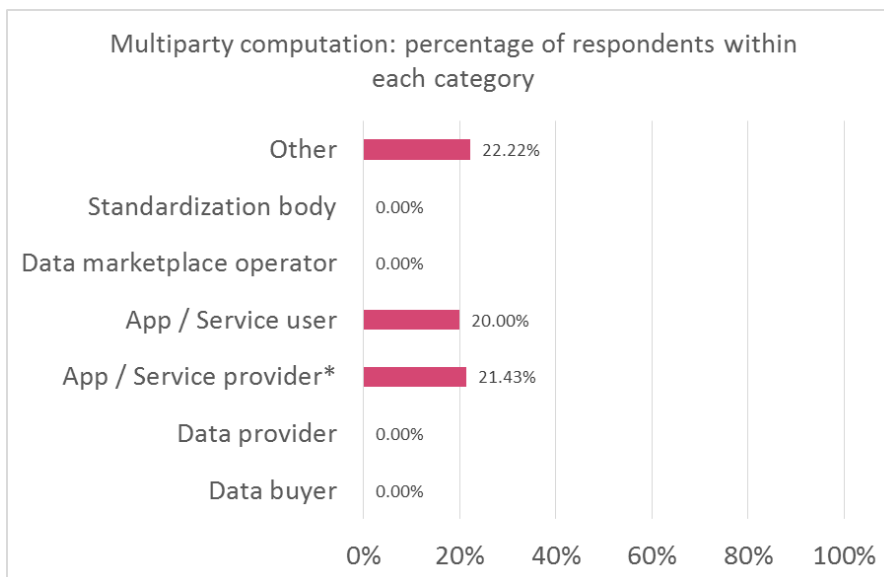


Figure 24: Multiparty computation (mandatory) – percentages of responses within organization roles' categories

The next question asked participants to identify the advantages of a data marketplace. Analysis of participants' responses identified the following main advantages:

- It consists of a **trusted** third party for data exchange.
- It provides an easy to find **common online place** to sell or buy data.
- It constitutes a **secure** medium for data sellers and buyers, ensuring also privacy, transparency and fairness.
- It guarantees dataset **anonymization**.
- It provides dataset **valuation**.
- It provides means for assessing datasets, for example through user reviews.
- It facilitates getting the data on a **self-service** and on-demand basis.
- It is a **reliable** place and a **safe** environment.
- It removes any physical location barriers.

- It ensures the **validity** of datasets.
- It provides a **variety** of datasets, including **multimarket** data.
- It provides **easier access** to data.
- It ensures **transparency** during data exchange processes.
- It **combats data monopoly** from the very big vendors.

Next, participants were asked to identify what should be avoided by a data marketplace. Analysis of participants' responses identified the following main concerns:

- Lack of clear rules for third party application inclusion or federation.
- Lack of matchmaking services, which would constitute it to a mere content delivery cloud service.
- Security risks: the marketplace should be 100% secure.
- Inclusion of datasets that are of poor quality or overpriced.
- Loss of privacy, by collecting and using data in uncontrolled ways.
- Usability risks: employing a complex structure and complex procedures for finding datasets or services.
- Insufficient paying and billing options.
- Sharing data without metadata support.
- Poor user help and documentation.

The last question pertained to the pricing models that participants preferred for a data marketplace. Results (Figure 25) revealed a clear preference over pay per use (69.57%), followed by options for free without Service Level Agreement (39.13%), fixed price subscription (30.43%), package (26.09%), and progressive price (13.04%). It can be easily assumed that each responder selected the option that best fits the needs of their organization; therefore in order to support the widest possible client base flexibility in options should be provided.

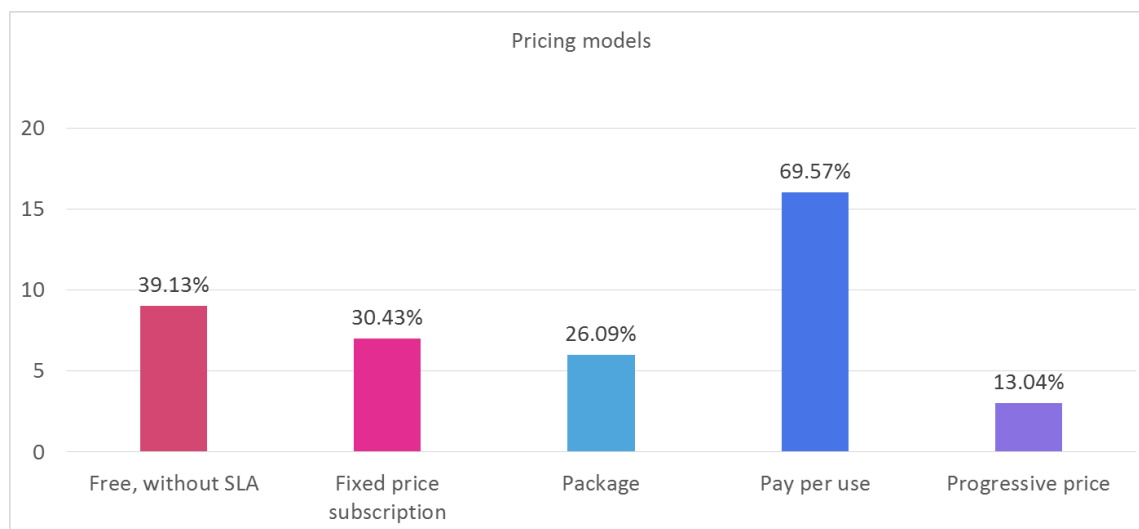


Figure 25: Pricing models

A chi square test of independence showed that there was no significant association between the type of organization and the preference over specific pricing models,  $\chi^2(8, 31) = 7.85, p = .44$ . In addition, there was no significant association between the role of an individual (i.e. domain expert, executive officer, operative officer, researcher) with the data marketplace functions selected,  $\chi^2(12, 31) = 11.28, p = .5$ . Last, a chi square test of independence indicated that the null hypothesis that the role of one's organization is independent from their preference regarding the data marketplace functions can be marginally rejected,  $\chi^2(24, 31) = 35.71, p = .05$ . Further analysis through chi-square tests for the effect of each variable on each pricing model revealed that two organizations groups exhibited strong, statistically significant, preference for the pay per use option, and

namely data buyers  $\chi^2(1, 31) = 4.28$ ,  $p = .03$ , and application / service providers  $\chi^2(1, 31) = 4.01$ ,  $p = .04$  (Figure 26).

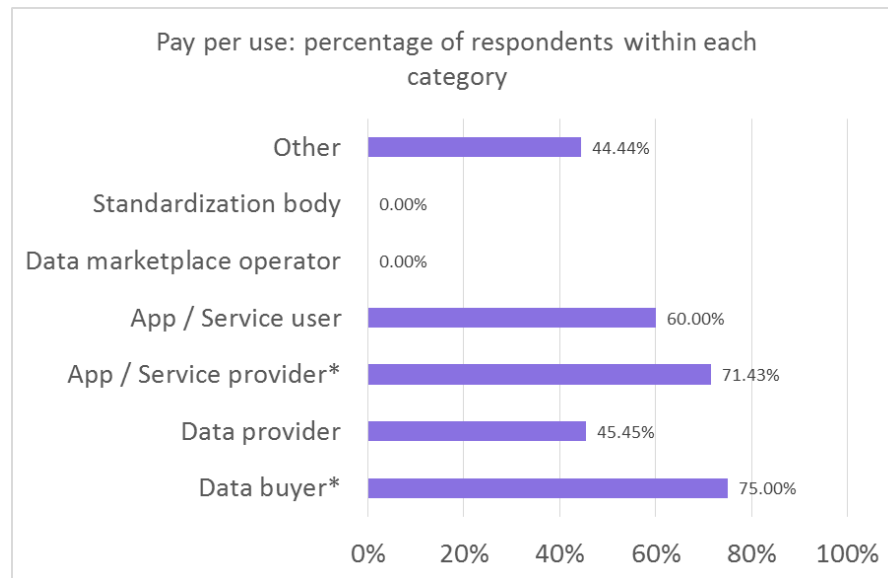


Figure 26: Pay per use – percentages of responses within organization roles' categories

#### 7.1.4 Requirements

This section consolidates the requirements, as they have been identified from the analysis of participants' responses to questionnaires. Although further insights and analysis of responses from electronic surveys is not possible, for each requirement the user group that asked for it is identified, and a connection with other requirements is provided whenever possible.

##### QR1 One-stop-shop online service for buying and selling data

###### Remarks

The requirement for **completing the entire process of buying and selling data** online emerged from both data buyers and sellers, but also by data application / service providers. Keywords such as electronic process, online process, one-stop-shop online service were used by multiple respondents.

##### QR2 The process should preserve privacy

###### Remarks

This concern was mainly pointed out by participants from data buyer organizations, but also from data application / service providers.

**QR3 Data buying and selling process should be standardized****Remarks**

This requirement was identified both by data buyers and data sellers. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace.

**QR4 The E2E platform implementation should provide easy access to the users****Remarks**

This requirement was identified by data buyers. Although through online surveys it is not possible to further elaborate on a finding / requirement, ease of access and ease of use were pointed out by various target groups when attempting to describe their requirements for the various phases of data selling or buying process.

**QR5 Data discovery should be based on the FAIR principles****Remarks**

This requirement was specified by data buyers. FAIR stands for Findable, Accessible, Interoperable, and Reusable. As this requirement advocates a principle in the field of data science, providing guidance on how to structure data and metadata, it is apparently related to a number of other requirements (e.g. QR4, QR7).

**QR6 The process should be accessible upon login on a centralized look up service****Remarks**

This requirement was identified by data buyers. It is also directly related to the requirement QR4, regarding ease of access to the service and QR38 referring to strong authentication mechanisms. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace.

**QR7 Searching for datasets should be easy****Remarks**

This requirement was identified by data buyers and data application / service providers, and is also related to QR5, QR8, and QR43. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, identifying that datasets search discovery services should be based on metadata, ontologies, etc.

**QR8 Searching through keyword should be supported****Remarks**

This requirement was identified by data buyers, data application / service providers and buyers, and is also related to QR5, QR7, and QR35. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, identifying that datasets search discovery services should be based on metadata, ontologies, etc.

**QR9 Prior to procuring a dataset, information on the dataset value should be provided****Remarks**

This requirement was identified by data buyers, highlighting the need for complete and accurate information before proceeding to a purchase action. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, regarding dataset valuation.

**QR10 Information about a dataset should include data description and owner****Remarks**

This requirement was identified by data buyers and data application / service providers, highlighting the need for complete and accurate information.

**QR11 The data should have straightforward mapping with the required business entities****Remarks**

This requirement was identified by data buyers, highlighting the need for appropriate data classification and matchmaking with end users' needs

**QR12 Processing of Personal data should be GDPR compliant****Remarks**

Ensure that personal data are processed in compliance to GDPR.

This requirement was specified by data buyers, and is also related to QR18 about GDPR compliance. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace.

**QR13 Efficient processes e.g. one-click payment****Remarks**

This requirement was identified by data buyers, who asked for a one-click process, highlighting the need for efficient interactions and processes.

**QR14 Provided payment is achieved, direct downloading should be supported****Remarks**

This requirement was identified by data buyers, further highlighting the requirement for one-stop-shop (QR1).

**QR15 Different subscription schemes should be supported****Remarks**

This requirement was identified by data buyers, data sellers and data application / service providers who also asked for annual license subscriptions. Analysis of the question regarding the pricing models of a marketplace revealed that flexibility in options should be provided, including for example fixed price subscriptions, packages, progressive prices, and pay per use.

**QR16 It should be possible to contact the data providers and choose the best solution / partnership****Remarks**

This requirement was identified by data buyers, and builds upon their need for trusted relationships with sellers and the corresponding matchmaking that should be supported by the platform to address their needs in the best possible manner.

**QR17 Smart contracts, that can be accomplished online, billing and transactions logs should be supported****Remarks**

This requirement was identified by data buyers, and is directly relevant with QR1 for a one-stop-shop online service. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace.

**QR18 The process should be GDPR compliant and approved****Remarks**

This requirement was identified by data sellers and by data application / service providers, and is also relevant to buyers' concern about data anonymization (QR12). This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, further identifying that different user role rights should be foreseen according to GDPR processes.

**QR19 It should be easy to describe the provided datasets through online metadata forms following standards****Remarks**

This requirement was identified by data sellers, further elaborating on their requirement for a standardized process (QR3). An indicative standard that was provided as an example is that of Dublin Core<sup>27</sup>. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, identifying that datasets search discovery services should be based on metadata, ontologies, etc.

**QR20 The process of data uploading should be easy (one-click)****Remarks**

This requirement was identified by data sellers. The example of contemporary online file transfer services was provided to illustrate their requirement for on-click uploading and confirmation upon successful accomplishment.

**QR21 The platform should provide powerful insights with respect to the provided data and services****Remarks**

This requirement was identified by data sellers, who requested powerful insights regarding the provided data (e.g. number of downloads per different time periods, geographical distribution, etc.)

**QR22 The platform should provide online financial management of data sells****Remarks**

**This requirement was identified by data sellers and is directly related to QR1 for one-stop-shop services.**

<sup>27</sup> <https://dublincore.org/>



**QR23 The platform should support direct requests from clients****Remarks**

This requirement was identified by data sellers, as well as by data application / service providers and buyers. It is also related to QR24, regarding the direct contact of data sellers with application or service providers.

**QR24 The platform should support targeted selling to specific clients****Remarks**

This requirement was identified by data sellers and is also related to QR23. The process as described involved selecting a specific application or service provider, announcing the data to them, and – upon agreement - contracting the provider.

**QR25 The platform should support classification of datasets according to their price (paid or free)****Remarks**

This requirement was identified by data sellers and data application / service provider, reflecting their need to easily identify datasets according to a number of attributes. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, identifying that datasets search discovery services should be based on metadata, ontologies, etc.

**QR26 The platform should support open data****Remarks**

This requirement was identified by data sellers and by data application / service providers, but also by data marketplace platform operators.

**QR27 The platform should provide connection to highly visited application marketplaces****Remarks**

This requirement was identified by data application / service providers, who explicitly mentioned as examples of marketplaces Google Play and Apple Store.

**QR28 Browsing content categories should be supported****Remarks**

This requirement was identified by data application / service providers. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, identified as a datasets catalogue functionality for datasets and thirdparty applications / services catalogue. In this case, clear rules for third party application inclusion or federation should stand.

**QR29 Each dataset should be characterized with the use of tags****Remarks**

This requirement was identified by data application / service providers. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, identifying that datasets search discovery services should be based on metadata, ontologies, etc.

**QR30 A dataset rating functionality should be supported****Remarks**

This requirement was identified by data application / service providers, reflecting their need for estimating the quality of datasets before purchase. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace.

**QR31 User comments on a dataset / service should be supported****Remarks**

This requirement was identified by data application / service providers and buyers, reflecting their need for estimating the quality of datasets before purchase.

**QR32 Information regarding the anonymization of a dataset should be provided****Remarks**

This requirement was identified by data application / service providers, reflecting their need for anonymized datasets. This requirement is also related to QR12.

**QR33 Before purchase, potential clients should be able to have a preview****Remarks**

This requirement was identified by data application / service providers and buyers, reflecting their need for estimating the quality of datasets before purchase.

**QR34 Networking between traders should be supported****Remarks**

This requirement was identified by data application / service providers, reflecting their need for building trusted relationships with their clients and dataset providers.

**QR35 Advanced searching through filters should be supported, including filters related to cost****Remarks**

This requirement was identified by data application / service buyers, as well as by data marketplace platform operators and standardization bodies. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, regarding dataset valuation.

**QR36 Information about an application / service should include any specifications****Remarks**

This requirement was identified by data application / service buyers.

**QR37 Comparisons between different services from different providers should be supported****Remarks**

This requirement was identified by data application / service buyers, reflecting their requirement to accurately identify the service/applications that is the most suitable for their needs.

**QR38 Strong authentication mechanisms should be provided****Remarks**

This requirement was identified by participants representing standardization bodies, and is also related to requirements QR4 and QR6.

**QR39 Intelligent matchmaking mechanisms should be provided****Remarks**

This requirement was identified by participants representing standardization bodies, and reflects their need for finding in an easy and accurate manner the data or applications / services needed.

**QR40 Brokerage****Remarks**

This requirement stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, and in particular datasets / applications trading service.

**QR41 De-anonymization risk analysis****Remarks**

This requirement stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace.

**QR42 Private set intersection****Remarks**

This requirement stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, regarding actions that can be performed on datasets from different parties, including multiparty computation.

**QR43 The platform should be usable****Remarks**

This requirement stems from the analysis of participants' responses to the questions regarding the risks that a data marketplace should avoid, but also from the individual answers of participants from all the organization categories, including keywords such as easy to use, easy to access, and easy to search. Overall, the platform should be usable, ensuring effective, efficient, and satisfactory user interactions. In this respect, clear help and documentation should also be provided.

## 7.2 Interviews

### 7.2.1 Method and Procedure

In the context of this task, a series of interviews have been conducted with stakeholders related to data buying and selling (providers and consumers). In detail, the TRUSTS partners have performed twelve one-to-one interviews with employees of a diversity of organizations that had direct or implicit interest in using a federated cloud-based European data marketplace. For the execution of the interviews, specific guidelines have been followed by the interviewers as described in detail in Annex II.

### 7.2.2 Participants

The following table summarizes the role of each interviewee in their organization as well as their relations regarding data or online service production / consumption. All interviews are included in Annex III.

Table 2: Interviewees consolidated information

Organization type	Role in the organization	Relation with datasets / services	Relation with TRUSTS
Data marketplace operator	IT project manager	<ul style="list-style-type: none"> <li>Implementation of a local data marketplace</li> </ul>	External
Telecommunications operator & content provider	Legal department	<ul style="list-style-type: none"> <li>A set of predefined contracts should be provided</li> <li>The local law for each federated node much apply</li> <li>Clear terms of use</li> <li>Regulations compliance</li> </ul>	Partner
Telecommunications operator & content provider	Data analyst	<ul style="list-style-type: none"> <li>Data exchange and warehousing</li> <li>Shared data repositories</li> <li>Analytics and marketing applications</li> </ul>	Partner
Telecommunications operator & content provider	Marketing strategy	<ul style="list-style-type: none"> <li>Internal historical aggregated and anonymized data from the marketing department, technical department, HR, finance.</li> <li>Projection analysis from respective consultants (e.g. Mason Analysis, IDC, etc.)</li> <li>Bespoke projection analysis and consultancy services.</li> </ul>	Partner
Telecommunications operator & content provider	Customer base management	<ul style="list-style-type: none"> <li>Access to new tools and methods for data analysis e.g. based on AI.</li> <li>Trend analytics/dataset on similar markets</li> <li>Possibility to predict behaviour using historical data (e.g. downgrades if a new</li> </ul>	Partner

		<ul style="list-style-type: none"> <li>pricing policy is enforced).</li> </ul>	
Legal corporate	Legal advisory; Financial compliance	<ul style="list-style-type: none"> <li>Data collection on a weekly basis, analysed by company's private means.</li> <li>Purchase &amp; sale of data in a weekly basis</li> <li>Purchase of data from open corporate services upon request of their clients.</li> </ul>	External
Audit and assurance	Audit, taxation and business advisory services	<ul style="list-style-type: none"> <li>Data from open corporate services providers are collected regularly and are analysed by the company's own tools.</li> <li>Purchase &amp; Sale of data on a monthly basis.</li> <li>Purchase of data from open corporate services, via an annual subscription (not a data marketplace).</li> </ul>	External
IT Development and business corporate services	IT project manager	<ul style="list-style-type: none"> <li>Purchase &amp; sale of data in a daily basis</li> <li>Possibility to predict behaviour using datasets.</li> <li>Daily collection of data from open corporate services providers and analysis by company's means.</li> <li>Upon client's request, access to the corporate provider, checks for abnormal behaviour and malicious patterns, alerts for any suspicious behaviour of client's data and analysis of results for the end-user.</li> </ul>	External
Bank	IT project manager in the field of banking applications	<ul style="list-style-type: none"> <li>Use of the available private data (e.g. customer profiling, analysing transaction patterns, past and immediate customer behaviour) to get real-time customer insights.</li> </ul>	External
Law firm	Dept service	<ul style="list-style-type: none"> <li>Analysing performance data and taking decisions on possible improvements and opportunities or new areas of business process improvements.</li> </ul>	External
Municipal service of European City	Data governance coordinator	<ul style="list-style-type: none"> <li>Governance of Municipal data such as stamp data and metrological data.</li> </ul>	External
Bank	Digital business/Customer Innovation	<ul style="list-style-type: none"> <li>Reinforcing and improving its internal digital services (such as the financial analysis and risk assessment)</li> <li>Identifying potential cooperation with 3rd party organizations.</li> </ul>	Partner

### 7.2.3 Requirements

Overall, all the interviewees expressed their eagerness for the TRUSTS results, since all agreed that getting access to a trusted data marketplace that will be able to accommodate a big number of data and services, respecting and conforming to the European laws and regulations about data privacy and management, would be a very useful tool in their daily work operations.

The findings that emerged by the interview analysis are summarized in the following requirements list. For each requirement a unique ID as well as explanation remarks is given.

#### IR1 Secure and legally compliant exchange of the datasets and services is required.

##### Remarks

The majority of the interviewees argued on the assurance that the TRUSTS platform should provide in respect to the integrity of the transactions performed between the producers and the consumers, as well as, the need for a legal and secure framework that will ensure the protection of the data that are made available in terms of privacy and infringement protection. It was a common suggestion from most of the participants that TRUSTS should respect and safeguard data access according to the international, European and national data protection laws and regulations (e.g., GDPR). Also compliance with ECB's regulations for financial data is required. Furthermore, many interviewees considered that this conformance capability should be exposed to the users through a comprehensive description of the terms of use. In addition local laws should apply to each federated node. A suggestion to facilitate business is to provide a set of predefined contracts.

#### IR2 Need for mechanisms that ensure the validity of the datasets and services onboarding process. Users' reputation schemes should also be supported as a protection measure.

##### Remarks

It was clear by the most of the interviewees that trust to the platform should be ensured by providing self-regulating mechanisms regarding on the one hand the validity and integrity of the onboarded datasets and services and on the other hand the validity of the providers. The existence of such mechanisms will act as key enablers for the buyers, to annotate and provide feedback that pertains to the quality of the datasets and services that they have bought, as a quality metric of the data and services a producer offers.

#### IR3 Due to the expected large number and vast diversity of the onboarding datasets and services, flexible pricing models, billing mechanisms and brokerage services should be provided. The integrity of the transactions between producers and consumers should be safeguarded through smart contracts, audit mechanisms and transaction logs, which must constitute an inherent part of the system.

##### Remarks

A common sense that was evident by all the participants is their need to use TRUSTS as a one-stop-shop service, through which they can find, bid for and buy available datasets and services. To that end they considered the existence of a billing system as well as brokerage services as granted. Another aspect that the interviewees considered as to be supported by TRUSTS is the implementation of flexible pricing models able

to be adapted according to the particular characteristics of the provided datasets and services. Finally, it was mentioned that it would be useful for the enterprises and companies to be able to create corporate accounts for their employees so that only one subscription/enrolment will be required.

#### **IR4 Standardised and easy processes should be provided for the onboarding of datasets and services.**

##### **Remarks**

Data providers agreed that the system should facilitate efficient mechanisms for the entry of datasets and services, which will not require much time for being accomplished and in parallel, help the users to avoid performing errors or entering misleading information during the onboarding process. For example, UI wizards can be provided for manual entry or application endpoints can be supported for the automatic and batch entry of datasets and / or services.

#### **IR5 Inherent data governance mechanisms should be provided. Continuous harvesting of new data for updating the semantic information of the onboarded datasets and services, are mandatory for effective cataloguing, products correlation and up to date descriptions.**

##### **Remarks**

Interestingly, besides the interviewee who works for the data market operator, many other participants stated that the comprehensive, well-structured and modelled meta-information of the provided datasets and services is mandatory for the system. By supporting a contemplating and effective data governance infrastructure, TRUSTS will be able to provide advanced functionalities, such as product-user matchmaking and recommendations, rather than basic search functionalities only, which are necessary and welcome as well. Thus, beyond keeping only basic information of the onboarded datasets and services, additional meta-data should be supported aligned with ontologies and taxonomies of different domains (science, industry, etc.). To that end, along with the appropriate profiling of the TRUSTS users, such a data governance scheme can lay the foundations for enabling opportunities for value added services that can be provided by the system (e.g., personalized catalogues that fit in users' needs, recommendation or matchmaking services, enhanced data combination, etc.). As such, data governance facilities are mentioned as necessary by the interviewees, noting that such functionality is missing from the tools that they use today. Furthermore, it was declared that the meta-data information, which is kept in the system, needs to be up to date through harvesting external sources on a frequent basis.

#### **IR6 Effective and secure user management should be employed.**

##### **Remarks**

Besides the profiling of users, datasets and services, one fundamental aspect that emerged by the interviews was the need for user management. In more details, within the TRUSTS environment, the users need to feel protected since they deem to make monetary transactions. To that end, strong authentication and authorization mechanisms should be provided, either to isolated users but also to enterprises and companies that have to give access to more than one of their employees. Furthermore, it was mentioned that each user should be aware of new products that fit in their need, in a timely manner, as well as be able



to announce to the marketplace needs for datasets and services.

#### **IR7 Use of intelligent profiling mechanisms for effective matchmaking and recommendation of datasets and services to the users.**

##### **Remarks**

Almost unanimously, the interviewees mentioned that a data marketplace, as TRUSTS, should provide anticipatory functionalities regarding the users' needs. In specific, it was explicitly mentioned that the system should provide smart profiling mechanisms that will exploit the data governance features of the platform, so as to be able (a) to match and recommend datasets and services as per users' needs; (b) to combine similar datasets for augmenting the available volume of data that fit in specific cases; and (c) to match datasets with services. Furthermore, it was mentioned that the system would facilitate the cooperation and interaction between data providers towards the provision of combined products that impose benefit to the consumers.

#### **IR8 Integration of Artificial Intelligence and specifically Machine Learning into the system so as advanced data analytics to be provided towards leveraging the capabilities of the services offered by the system (e.g., recommendations, matchmaking, etc.)**

##### **Remarks**

Indisputably, nowadays AI and ML constitutes important tools for many businesses today towards improving their performance and effectiveness. To that end, participants mentioned that such mechanisms would be very useful to be provided by the system, thus leveraging its capabilities in many different fields such as pattern identification applications, test/simulate recommendation engines, better detection and accuracy, data mining and predictive analytics.

#### **IR9 Inherent protection of private datasets should be provided.**

##### **Remarks**

The majority of the interviewees need to gain access to private data, which many times might originate from the processing of sensitive / personal data. Thus, the protection of such datasets through anonymization mechanisms that will be able to be applied on the datasets during their onboarding process and before they are published, is more than necessary according to the participants' opinion. Furthermore, some of the interviewees stated that it would be very useful if de-anonymization risk assessment could be provided as a protection measure for the private anonymized data that the TRUSTS users' aim to publish. Finally, private datasets intersection, through cryptographic techniques that allows two parties to combine data in an encrypted manner in order to be able to compute their intersection (all relevant protection approaches can be applied e.g. PSI/MPC, masking common parameters to datasets that are used for correlation, etc.), is also very welcome.

**IR10 Data marketplace should be easy and friendly to use, leveraging productivity and decreasing operational costs through an enriched cost-effective functionality.**

#### **Remarks**

A general comment that emerged by the majority of the participants, was the need for an easy and friendly to use data marketplace, which is able to provide intuitive and comprehensive functionality in the most productive way. This approach will conclude to the mitigation of the companies' operational costs in their quest of selling or buying data and services.

## **8 Detailed Use Case Definition and Target Business & Technical KPIs**

Key role in the requirements collection process and the KPIs definition for the project trials plays the detailed definition of the three real life TRUSTS use cases in terms of stakeholders, required data marketplace functionality and services, targets and high level scenarios.

### **8.1 Use Case 1 - The (AML) compliance use case**

A European Data Marketplace as it is TRUSTS represents an opportunity to the data market of data sharing securely and GDPR compliant such as trade data for AML purposes, thus maximizing its operational effectiveness.

The ambition of EBOS, FNET and InBestMe is to classify not only the business convenience but also the technological opportunities that are derived from the TRUSTS data marketplace. Furthermore, the aforementioned companies based on their experiences and their business perception aim at contributing to the requirements and specifications of a data marketplace so as to be able to enhance the efficiency and the sustainability of Anti-Money Laundering services that the TRUSTS data marketplace may offer to the market.

The lack of a consolidated and widely viable data marketplace, secure and GDPR compliant adequate to benefit various business collaborations in the framework of AML services enhanced with Artificial Intelligence consist a necessity to the data market. Such marketplace collaboration will be a benefit for the whole economy since innovative procedures and productions with added value will be inaugurated into the market.

The aforementioned will be tested for their usability, efficiency and operation completeness through UC1 trials. The envisaged marketplace operations and the respective marketplace services are illustrated in Figure 27:

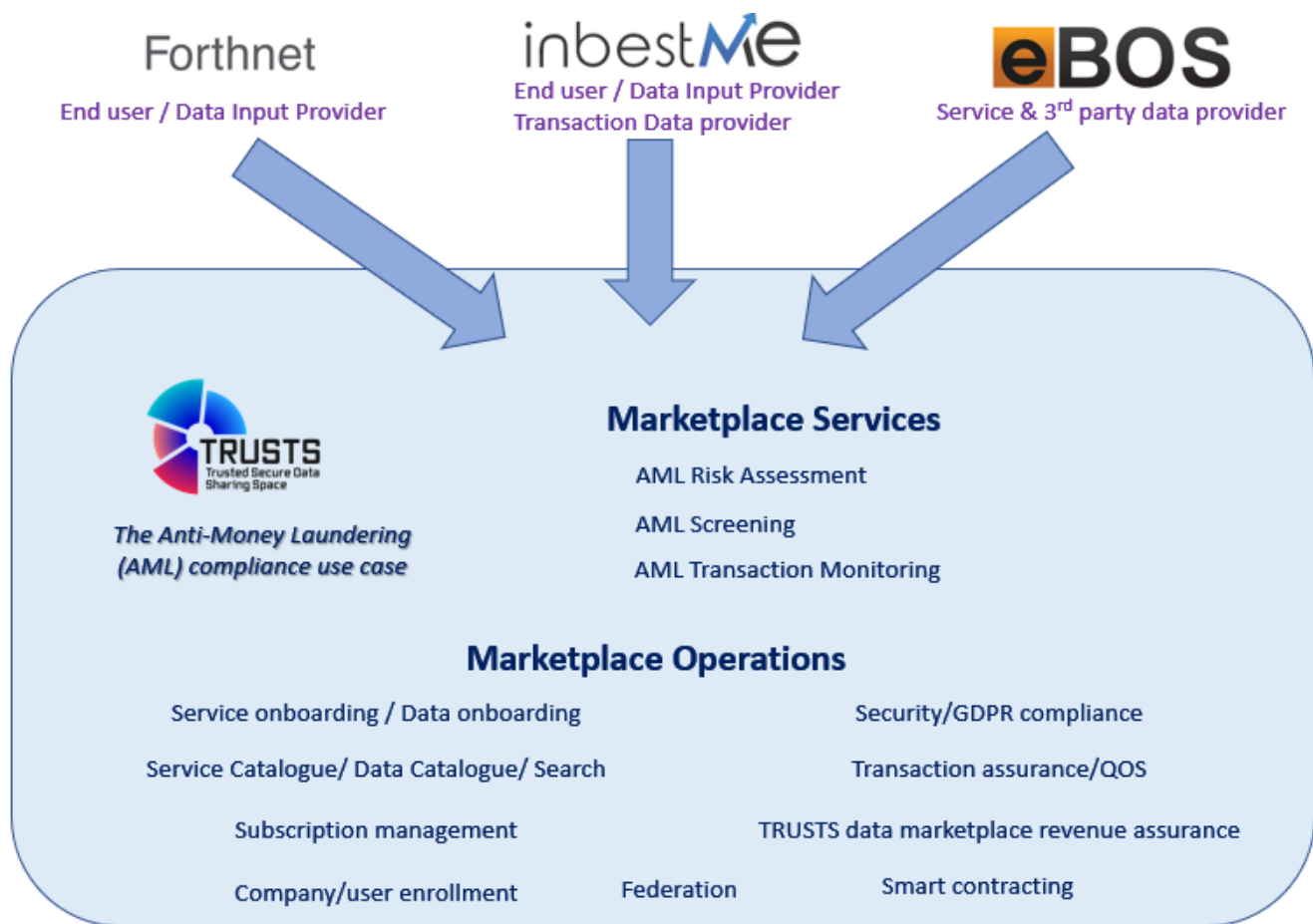


Figure 27: The AML process through the TRUSTS data marketplace

### 8.1.1 UC1 data marketplace requirements and trial definition

UC1 envisages TRUSTS as an end to end environment able to offer quality process and services to meet the UC1 trial needs.

Specifically, TRUSTS should have a coherence set of operational functionalities providing the ability to offer business services and datasets:

#### TRUSTS operational functions:

In order to sustain its operation TRUSTS should support at least the following operational functionality:

- **Service on-boarding process:** AML services can be onboarded to the TRUSTS data marketplace via respective functionality offered by the marketplace for uploading the service solution files. The on-boarding (uploading) process should be secured via a login or/and token authorization functionality. This on-boarding process should also include a testing process (so that the service can be functional).

- **Data/Meta data on-boarding process:** similar to the above, TRUSTS data marketplace shall allow the on-boarding of data, by choosing the appropriate service to be used for. The process again should be secured as above.
- **Service/data catalogue and search:** Both services and data provided in TRUSTS data marketplace, after on-boarding, shall follow a process of approval to include them in the marketplace catalogue and search functionality throughout all the federated nodes. That means the user can search for them even directly even by a number of keywords in the Platform's search engine.
- **Subscription management / Company and user enrolment:** TRUSTS data marketplace as mentioned above should be secured via a login or/and token authorization functionality allowing the subscription and user enrolment of companies and with specific roles within the subscribed companies' users/employees.
- **Security and GDPR compliant procedures:** securing users' data on subscription and during the marketplace usage (purchase, smart contract, billing etc.) following GDPR rules. Accept policies option should also be provided to the user (during subscription, login, and purchase).
- **Purchase Process assurance/QoS:** Transactions in respect to the purchase of a service should be transparent and secure following the global transaction policies and GDPR. Transactions logs should also be kept by the marketplace to be used in case of a transaction issue occurred. The option to the user to validate the service (after purchase it and use it) should also be provided.
- **Revenue assurance:** Trusts data marketplace should provide a mechanism to assure the revenue of the AML service provider for each purchase based on the Billing, Compensation of involved parties in the value chain.

### UC1 specific services:

The required TRUSTS services for UC1 trials as per Figure 27 above are:

- **AML Screening:** This is the first process that a customer (i.e.: fiduciary, law firm, etc.) must do before the onboarding of its clients (or potential clients) either those are physical persons or legal entities. The core purpose of customer screening is to add to the risk picture of the customers (or potential customers) and, specifically, to identify if they are for example subject to international sanctions. It delivers a sophisticated screening solution capable of screening customers against Politically Exposed People (PEP) lists, Sanctions lists, Adverse / Negative Media reports, etc.

This service will use the input data (KYC information<sup>28</sup>) from the end user in order to check an entity (a person / company), in reference with the data provided by RDC<sup>29</sup>, (3<sup>rd</sup> party data provider). RDC provides various datasets (i.e. PEP lists, Sanction lists, Adverse Media). As a result, this service will provide a report with the profile of the checked entity *if for example a person is a Political Exposed – PEP, or a person it is found in Sanction list. As an example, in case of an entity appears in Sanctions List, then the end-user must not collaborate with this client anymore or must not proceed with any potential collaboration).*

- **AML Risk Assessment:** Normally, this service follows the screening process. Risk Assessment provides a wide view of customers' relationships, rates their risks based on a dynamic rules-based engine, monitors activities, detects, investigates and documents suspicious cases. This service will use the input data that the end-user will provide for ex: business activities of the customer, country of incorporation, destination of funds, country of origin of funds, etc. as well as the screening report for the

<sup>28</sup> For legal entities (company): name of the company, jurisdiction and registration number. For physical person: name, surname, nationality

<sup>29</sup> <https://rdc.com/>

aforementioned entity (as this can be extracted from the AML Screening service) and a ruled based questionnaire as per ESA guidelines<sup>30</sup>. As a result, this service will categorize the entity *as a low/medium/high risk*.

- **AML Transaction Monitoring:** Transaction monitoring is usually based on financial transactions and it is a vital part of preventing money laundering and terrorist financing. This service will use the input data from the end user, acting on behalf of its customer by uploading in the system documents such as agreements and the relevant transaction information. Based on engine rules this service will check if the transactions are in conflict with the input data (i.e.: upper the limit of the daily basis), if are valid, relevant to the agreement, etc. and identify any possible frauds or anomalies on those transactions. All checks performed by the above-mentioned provided AML services along with the validity of the results are highly dependent from the input data provided by the end user. No further cross checks are performed with any additional fiduciary or financial institutions.

The above-mentioned services will be built on the TRUSTS federated infrastructure, employing the necessary components so as to enable the secure data exchange, to safeguard the private information under a technical and legal perspective, but also preserve the capability to deliver reliable results and insights. All input data and AML check results, after UC trial (and service execution) will be anonymized and stored in a database located to TRUSTS platform, and will be used for the purposes of the project and the UC1, to serve the Machine Learning and AI functionality (aims to be implemented in UC1). Those data will be kept only for the duration of the project.

#### UC1 high-level trial scenarios:

In order to test the end-to-end service and provide valuable feedback towards improving both technological and business aspects of the TRUSTS data marketplace, the following high level trial scenarios are envisaged:

Table 3: UC1 high level trial scenarios

UC1 Trial Scenarios	Description
<b>Service and testing data onboarding</b>	On boarding of AML services and data to be tested for verifying the successful onboarding (incl. the smart contract).
<b>Companies subscription</b>	FNET & InBestMe subscription (selection of plan, subscription, signing the contract/smart contract, companies representative's definition and roles, logs existence).
<b>Service catalogue usage</b>	Search by the end user in service catalogue either by key words or directly for the adequate AML service in all federated nodes

<sup>30</sup> [https://esas-joint-committee.europa.eu/Publications/Guidelines/Guidelines%20on%20Risk%20Factors\\_EN\\_04-01-2018.pdf](https://esas-joint-committee.europa.eu/Publications/Guidelines/Guidelines%20on%20Risk%20Factors_EN_04-01-2018.pdf)

<b>Contract fulfilment</b>	Ensure smart contract fulfilment
<b>Billing and Payment</b>	Completion of the billing and payment cycles. Marketplace keeps tracking (logs) of the transactions.
<b>Data Onboarding</b>	Onboarding of InbestMe input data after choosing the adequate AML Service.
<b>Service usage</b>	Schedule service usage, use the service, and evaluate the service results.
<b>Service quality evaluation</b>	Collect users' evaluation and if needed to improve operations.

### 8.1.2 UC1 roles

In the following section, the role of each partner is detailed along with the interaction of the partners within the UC. The diagram that follows also presents those roles in a more graphical representation.

- **EBOS** will act as a service provider by on-boarding to the TRUSTS data marketplace the WiseBos AML services. (Risk Assessment, Screening and Transaction Monitoring).
- **EBOS** acts also as a data provider, by utilizing data from RDC (3<sup>rd</sup> party data provider). To do that, EBOS has a signed agreement in order to access those data. Those data is related to PEP lists, Sanction lists, Adverse Media and all of them are considered as private data (since a subscription to access them is required). The input data (provided by the end-users) includes physical and legal entities information (KYC, etc.), will be evaluated/checked for any AML suspicious activities. Those checks will be based on provided RDC data.
- **InBestMe** will act as data provider by on-boarding to the TRUSTS marketplace financial transaction data. This type of data along with the input data that the end-user will provide will be correlated.
- **InBestMe** will also act as an end-user. As part of its role InBestMe will provide input data about physical and legal entities information (KYC, etc.). AML checks will be performed on those provided input data. InBestMe will search for the AML services either directly or with key words through the search engine. InBestMe will proceed with smart contract, billing and then will be able to use the adequate AML services through the TRUSTS data marketplace.
- **FNET** will act as an end-user. As part of its role InBestMe will provide input data about physical and legal entities information (KYC, etc.). AML checks will be performed on those provided input data. FNET will search for the AML services either directly or with key words through the search engine. FNET will proceed with smart contract, billing and then will be able to use the adequate AML services through the TRUSTS data marketplace.
- **TRUSTS** will act as user administrator allowing the subscription and user enrolment of companies and with specific roles within the subscribed companies' users/employees.
- **TRUSTS** will also act as service administrator so as to accept the adequate services.

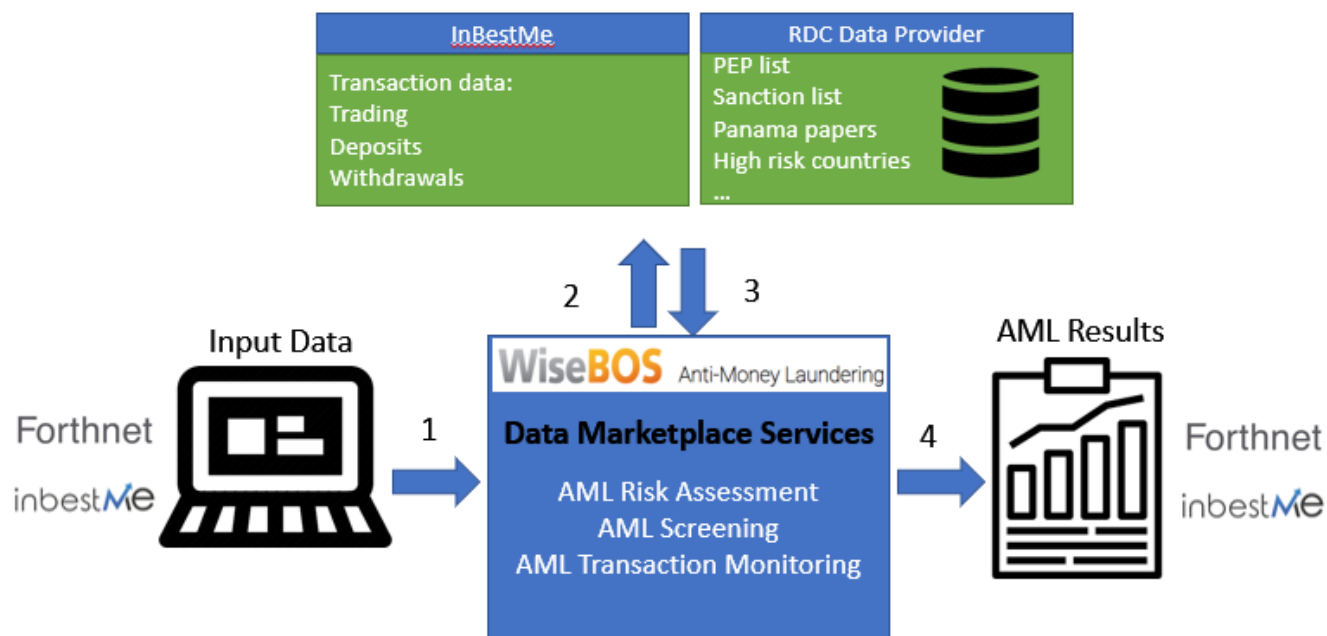


Figure 28: UC1 Partners Roles

The implementation of AML Services along with a correlation with data providers (RDC & InBestMe) will increase the results quality and fulfil the AML process so as to send accurate results to the end users (FNET & InBestMe) through the TRUSTS data marketplace.

### 8.1.3 UC1 KPIs

UC1 trials will be evaluated using the T2.3 methodology.

In particular UC1 will set the following KPIs:

Table 4: UC1 performance and process KPIs

Process	KPI
Service and test data onboarding	<p><b>Description:</b></p> <p>End to end service &amp; testing data onboarding process to be fulfilled.</p> <p><b>KPI:</b></p> <p>At least <b>two versions (w/o AI and with A/I) of the AML applications</b> are successfully on-boarded on TRUSTS nodes.</p>
Companies subscription	<p><b>Description:</b></p>

	<p>User friendliness, clear processes, ability to verify and modify, logs existence</p> <p><b>KPI:</b></p> <p>Successful subscription of EBOS, FNET and InbestMe. Successful definition of roles. Successful enrolment of FNET and InbestME, representatives.</p>
<b>Service catalogue usage</b>	<p><b>Description:</b></p> <p>Search in catalogue using keywords throughout all federated nodes. (search data &amp; service)</p> <p><b>KPIs:</b></p> <ul style="list-style-type: none"> <li>• Return adequate response in &lt;1sec.</li> <li>• User task success &gt; 90%</li> <li>• User satisfaction, SUS<sup>31</sup> score &gt; 70</li> </ul>
<b>Data onboarding</b>	<p><b>Description:</b></p> <p>On boarding of InbestMe data/metadata by choosing the adequate AML Service</p> <p><b>KPI:</b></p> <p>Successful onboarding of data as per the specific file type</p>
<b>Service usage</b>	<p><b>Description:</b></p> <p>Well-structured, defined modules deployment, if necessary, process.</p> <p><b>KPIs:</b></p> <ul style="list-style-type: none"> <li>• Customer loyalty NPS<sup>32</sup> &gt; 8 [0-10]</li> <li>• User satisfaction, SUS score &gt; 70</li> <li>• Detailed results analysis, SUS score &gt; 85</li> <li>• Service excellence, SUS score &gt; 80</li> </ul>
<b>Contract fulfilment, service performance tracking, quality evaluation</b>	<p><b>Description:</b></p> <p>Contract fulfilment, transaction logs existence, user evaluation existence, process to evaluate complete process by the TRUSTS operations in order to improve performance existence.</p> <p><b>KPIs:</b></p>

<sup>31</sup> <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

<sup>32</sup> <https://www.netpromoter.com/know/>



	<ul style="list-style-type: none"> <li>• At least 2 contracts are fulfilled.</li> <li>• Operation completeness, SUS score &gt; 80</li> </ul>
--	--

### 8.1.4 UC1 longer terms business KPIs

The KPIs defined in DoA and the process to meet them is outlined below:

Table 5: UC1 longer term business KPIs

KPI	Baseline Value	Target Value (M36)	Process to meet the target KPIs
Number of alerts per scenario	Number of alerts per scenario issued by WiseBOS ERP solution	Decreased by 50% from baseline	Perform adequate number of trials
Detection accuracy	Detection accuracy from WiseBOS ERP solution	Increased by 50% from baseline	Perform adequate number of trials
Number of false positives	Number of false positives flagged by WiseBOS ERP solution	Reduced by 30% from baseline	Perform adequate number of trials
Number of false negatives	Number of false negatives flagged by WiseBOS ERP solution	Reduced by 30% from baseline	Perform adequate number of trials
SAR capture	70%	>95%	Perform adequate number of trials
Losses due to fraud	As per self-assessment from end-users	Reduced by 30% from baseline	Perform adequate number of trials
Number of data providers interacting with the Platform	2 at the start of the use case	Minimum 10 by M36 (+400%)	In order to achieve this the project needs to involve additional data providers using dissemination activities
Number of end-users interacting with the Platform	1 at the start of the use case	Minimum 10 by M36 (+400%)	In order to achieve this the project needs to involve additional data providers using dissemination activities

## 8.2 Use Case 2 - Agile Marketing through Data Correlation

TRUSTS will constitute a European Digital Marketplace platform, which will enable FNET and PB to increase digital transformation and respective entrepreneurship activities toward being pioneers in the Greek Telecom and Banking sectors.

The aim of both FNET and PB is to evaluate business and technological opportunities that the TRUSTS data marketplace may offer. In addition, both enterprises aim to contribute to the requirements and specifications of the data marketplace using experience and company vision in order to increase effectiveness of the offered services.

The challenging envisioned business process of correlating external data sources in a GDPR and other respective regulations compatible manner e.g. anonymised and aggregated CRM data of FNET and PB, has been chosen as a base evaluation scenario. Current practices e.g. absence of a unified and commonly acceptable technological and business framework able to assist such business collaboration, make it difficult to explore such business opportunities since all respective negotiations have to start each time from the beginning. Nevertheless, both FNET and PB understand that such collaboration will be beneficial for both the companies and the clientele since it will lead to better products targeting real subscriber/client needs. The whole economy will be benefited as well since innovative process and product production value chains will be established. Such innovative processes will be tested through UC2 trials for their user friendliness, completeness and business effectiveness.

The envisaged operation and the respective trial definition are illustrated in Figure 29.

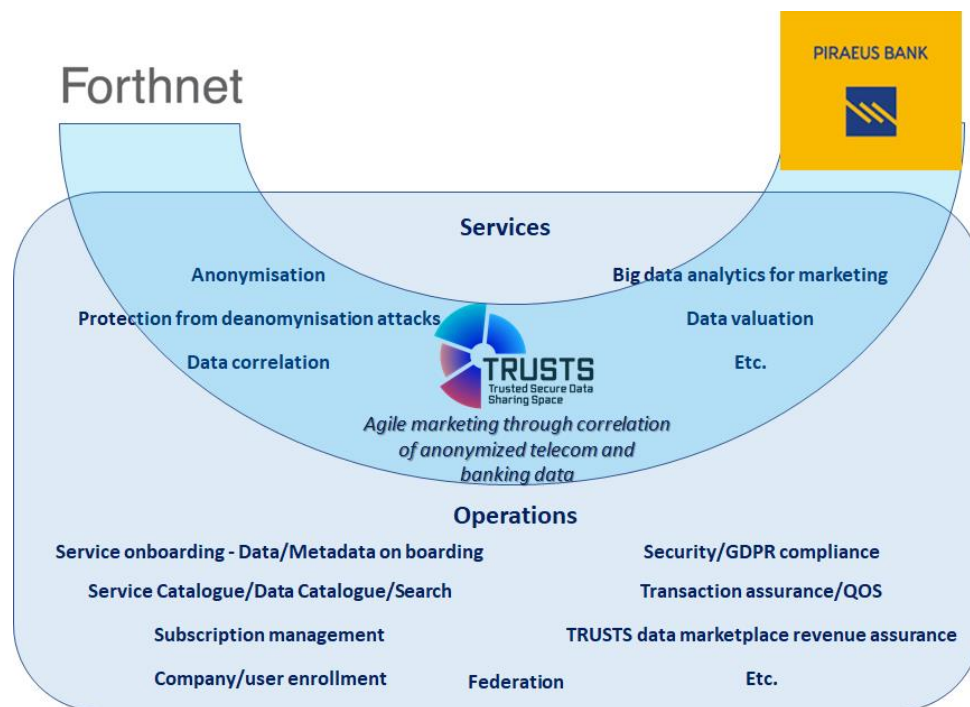


Figure 29: UC2 agile marketing trial architecture

### 8.2.1 UC2 data marketplace requirements and trial definition

UC2 envisages TRUSTS as an end to end environment able to offer quality process and services to meet the UC2 trial needs.

Specifically TRUSTS should have a coherence set of operational functionalities providing the ability to offer business services and datasets:

#### **TRUSTS operational functions:**

In order to sustain its operation TRUSTS should support at least the following operational functionality:

- Process to service on-boarding: Services can either be embedded in the TRUSTS data marketplace offered through a respective API. The on-boarding process should include testing of technical, performance and security aspects, smart contract establishment and inclusion in the services catalogues and search engine.
- Data/Meta data on-boarding: Data descriptors, lifecycle, smart contract, inclusion of the catalogues and search engine.
- Subscription management and contracting with client companies/users subscription with specific roles within the subscribed companies.
- Privacy and GDPR processes
- Billing, compensation of involved parties in the value chain, Revenue assurance
- Transaction logs
- Catalogue/Beyond the state-of-the-art search engine (e.g. matchmaking, recommendation, etc.) for services, data/metadata, etc.
- Federation/Transparency in service/data/subscription/service catalogue

#### **UC2 specific services:**

The required TRUSTS services for UC2 trials are:

- Anonymization: Nice to have. Anonymization in the trial will be done prior to data entering TRUSTS but is a feature that TRUSTS must have.
- Deanonimization risk analysis: All data must be checked for potential risks
- MPC/PSI: secure intersection of data without having access to the other party data
- Reporting: Transaction logging compliant to GDPR
- Operational/Subscription/Federation/Quality services

The above mentioned services will be built on the TRUSTS federated infrastructure, employing the necessary components so as to enable the secure data exchange, to safeguard the private information under a technical and legal perspective, but also preserve the capability to deliver reliable results and insights. Furthermore, these services will be made available to FNET and PB through TRUSTS to participate in the overall assessment of the platform with regard to data and services discovery and brokerage.

#### **UC2 high level trial scenarios:**

In order to test the end-to-end service and provide valuable feedback towards improving both technological and business aspects of the TRUSTS data marketplace, trials will follow the subsequent phases:

Table 6: UC2 high level trial scenarios

UC2 Trial Scenarios	Description
<b>Service Onboarding</b>	On boarding of MPC/PSI (onboarding, smart contract, inclusion to the service catalogue, quality test). Federation issues should be tested e.g. service onboarding in different federated nodes.
<b>Companies subscription</b>	FNET and PB subscription (selection of plan, subscription, signing the contract/smart contract, companies' representatives definition and roles). Federation issues should be tested e.g. companies subscribed in different federated nodes.
<b>Service catalogue usage</b>	Search in service catalogue by FNET and PB for the adequate MPC/PSI, deanonymisation risk analysis, etc. services. Federation issues should be tested e.g. transparently searching to all federated nodes.
<b>Service usage</b>	Schedule service usage (MPC, PSI, De-anonymisation risk analysis, end to end TRUSTS service), deploy any necessary modules, use the service, evaluate the outcome
<b>Contract fulfilment, service performance tracking, quality evaluation</b>	Ensure smart contract fulfilment, evaluate transaction logs, collect users' evaluation, improve operations if necessary.

### 8.2.2 UC2 roles

Figure 30 depicts the conceptual architecture of the TRUSTS' second use case, highlighting the contribution of each participating partner as well as the actions that need to be performed during the use case execution.

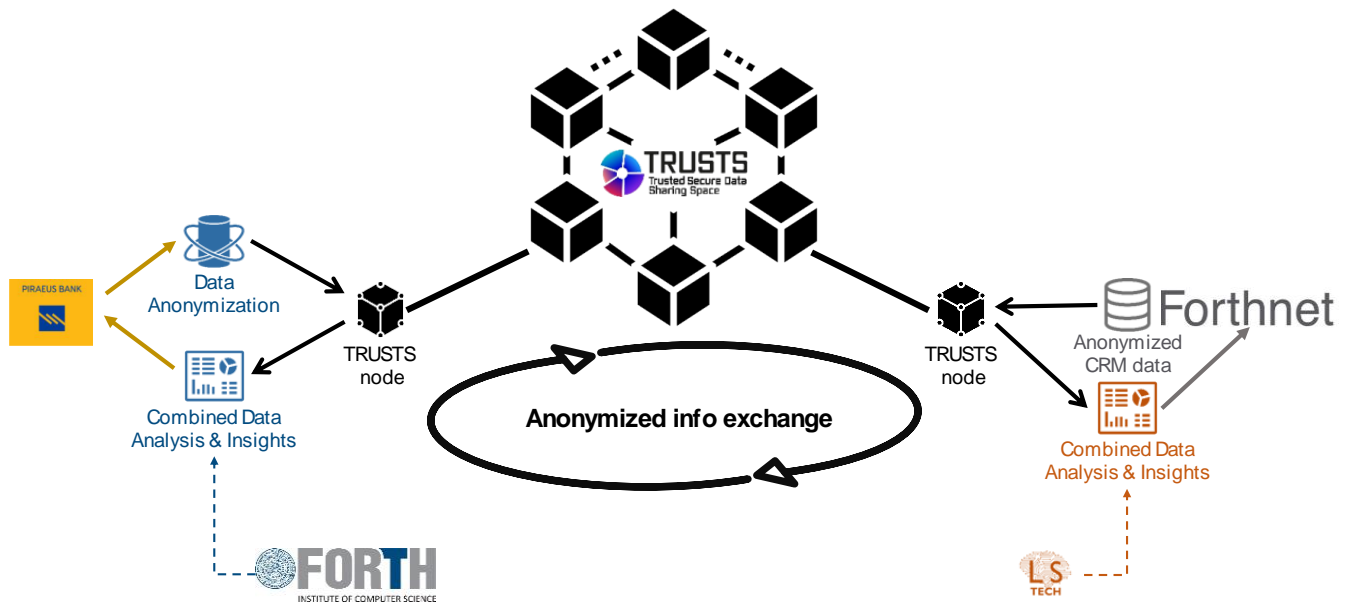


Figure 30: High-level conceptual architecture of the “Agile marketing activities through correlation of anonymized telecom and banking” use case

Specifically:

- TRUSTS data marketplace operators which will provide all the necessary services with adequate quality.
- FNET will provide anonymized CRM data and define the detailed business scenario. FNET will analyse the trial outcome
- LST will install if needed and use TRUSTS services to intersect FNET and PB datasets.
- PB will offer CRM data
- FORTH will install if needed the required TRUSTS to intersect FNET and PB datasets. In addition FORTH will provide Smart dashboards and big-data analytics and Customers’ economic behaviour insights.

### 8.2.3 UC2 KPIs

UC2 trials will be evaluated using the T2.3 methodology.

In particular UC2 will set the following KPIs:

Table 7: UC2 performance and process KPIs

Process	KPI
<b>Service Onboarding</b>	<p><b>Description:</b></p> <p>End to end service onboarding process to be fulfilled.</p> <p><b>KPI:</b></p> <p>At least <b>PSI/MPC, deanonymisation risks analysis applications</b> are successfully on-boarded on TRUSTS nodes.</p>

<b>Companies subscription</b>	<p><b>Description:</b></p> <p>User friendliness, Clear processes, ability to verify and modify, logs existence</p> <p><b>KPI:</b></p> <p>Successful subscription of FNET, PB, FORTH and LST. Successful definition of roles. Successful enrolment of FNET, PB, FORTH and LST representatives.</p>
<b>Service catalogue usage</b>	<p><b>Description:</b></p> <p>Search in service catalogue using key words across all federated nodes.</p> <p><b>KPIs:</b></p> <ul style="list-style-type: none"> <li>• Return adequate response in &lt;1sec.</li> <li>• User task success &gt; 90%</li> <li>• User satisfaction, SUS<sup>33</sup> score &gt; 70</li> </ul>
<b>Service usage</b>	<p><b>Description:</b></p> <p>Well defined applications (i.e. MPC, PSI, deanonymisation risk analysis, TRUSTS end to end service, etc.), modules deployment, if necessary, process.</p> <p><b>KPIs:</b></p> <ul style="list-style-type: none"> <li>• Customer loyalty NPS<sup>34</sup> &gt; 8 [0-10]</li> <li>• User satisfaction, SUS score &gt; 70</li> </ul>
<b>Contract fulfilment, service performance tracking, quality evaluation</b>	<p><b>Description:</b></p> <p>Contract fulfilment, transaction logs existence, user evaluation existence, process to evaluate complete process by the TRUSTS operations in order to improve performance existence.</p> <p><b>KPIs:</b></p> <p>At least 3 contracts fulfilment.</p>

#### 8.2.4 UC2 longer term business KPIs

The KPIs defined in DoA and the process to meet them is outlined below:

<sup>33</sup> <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

<sup>34</sup> <https://www.netpromoter.com/known/>

Table 8: UC2 longer term business KPIs

Key performance Indicator	Baseline value	Target value (M36)	Process to meet the target KPIs
Number of target marketing analysis	2 per month	>10 per month	Perform adequate number of trials
Data readiness for correlation	Low (1 week for data to become ready)	High (1 day for data to become ready)	UC data providers should provide adequate datasets for the trials
Data valuations	2 per month	>10 per month	Perform adequate number of trials
Data anonymisations/deanonymisations	<1 per month	>10 per month	UC data providers should provide adequate datasets for the trials
Number of data providers interacting with the Platform	2	>10	In order to achieve this the project needs to involve additional data providers using dissemination activities
Number of end-users interacting with the Platform	2	>10	In order to achieve this the project needs to involve additional data providers using dissemination activities

### 8.3 Use Case 3 - The data acquisition to improve customer support services use case

The TRUSTS Data Marketplace vision is to **create an out-of-the-box analytics solution for the anonymization and visualisation of big data**, specifically to **advance new ways of human-computer interaction currently in their infancy, e.g. chatbots that can act as automated assistants to allow customers to converse about the management of their debt at their own pace and with a personalized experience, through the integration of Big Data.**

The integration of cognitive computing will transform the industry in a variety of ways, from stimulating new ways to interact with customers (including some that, paradoxically, feel more human) to automating recurring tasks or helping detect patterns in data. TRUSTS will combine advanced natural language processing capabilities with the insights of Big Data as the basis for the development of tailored wealth management services. **Big Data** plays a significant role by enabling sophisticated **personalization** services that allow to classify each customer over a set of “**customer types**” based on their activity (thus giving the bot the ability to offer options for debt management tailored per customer case), and also by enabling a **personalized interaction** with respect to the tone and feel of the conversation by using sophisticated real-time metrics (e.g.,

emotion detection), making the experience as pleasant as possible for the customer. The main challenges can be summarized here:

1. Use of artificial intelligence to adapt human-computer interaction (HCI)/services to customers demands and needs.
2. Personalized approach: Demonstrate the impact of being aware of and responsive to users' context and human emotions.
3. Demonstrate the impact of more natural ways of communications (natural language, speech, facial expressions).
4. Seamless integrated data interconnection among a variety of end-user devices and applications.
5. Demonstrate in real field the outcomes of more natural online interaction mechanism to demonstrate the expected impact.

The purpose of this **demonstrator** is the development of a ground-breaking offering in the field of debt collections – that is a fully automated debt collections call centre, leveraging the power of the TRUSTS Platform. The idea is that through **enhanced analytics**, **AI** and the integration of **bots**, a **customer** will be able to run a full operation around **debt collection** without needing to employ agents to follow-up with customers.

The piloting activity around this will be performed in a small scale and with a controlled set of data and customer entries. **Relational Romania** in collaboration with **Debt Servicer X** will generate anonymised benchmark datasets using data management procedures that include **anonymization** and **cryptographic protocols** that will be set up to transmit all the data. Relational Romania through the TRUSTS Platform will improve its **chatbot's** application for the Debt Collection System and to provide the following two pilot applications through using data from **Debt Servicer X**:

1. **Debt Servicer X**: piloting debt collection with no employees. Everything is done through **human-computer** interaction. For the debt collection call centres (large in terms of personnel and cost), the bots could eventually enable the “agent-less collection centre”.
2. **Incorporation of a chat-bot** to act as an automated **assistant** that allows customers to converse about the management of their debt at their own pace and with a personalized experience
3. **Piloting chatbot** at website and mobile application users via both secure and unsecure channels respectively. The aim of the chatbot would be to market new offerings and increase usage ratios, **offering proper type of new offerings based on transaction history of the customer**.
4. **REL** will have access to large, complex and realistic **data** of consumer data from **Debt Servicer X** that will be combined with the existing data that the company has from the Point of Sales with Business Industry, transactions and consumer social activity. The new communication channel (hereafter referred as Chat Service), will be uploaded to the **TRUSTS Data Marketplace**, will be used by **FNET**.

### UC3 Process:

Relational (REL), in collaboration with the Creditor/Banking Organization will perform data anonymization/masking techniques that will protect and anonymize input data owned by the Creditor, just before any interaction with TRUSTS data marketplace to prevent any issues regarding data privacy.

According to our existing preliminary analysis, the AI models can be trained on anonymized data and we expect to deliver the KPIs without access to personal data.

REL will provide AI models related to debt collection that will get trained and will evolve using all input data available on the marketplace, related to Use Case 3.



In the context of UC3 scope, REL will **develop the following two services:**

- Automated decision making on the next-best-action for curing the debt collection issue using AI models trained from the data retrieved from the Marketplace.
- Automate communication with debtors using chat-bot technologies. Bot interactions again will utilize both the behavioural history but also would need a history of chat/messaging interactions.

### **8.3.1 UC3 data marketplace requirements and trials definition**

#### **TRUSTS operational functions/operations:**

In order to sustain its operation TRUSTS should support the following operational functionality:

- On-boarding of external data
- Services on-boarding
- Metadata discovery (catalogue) and maintenance (descriptions, tags etc.)
- Service usage and billing
- GDPR related certifications
- Logging and auditing

#### **Use Case 3 specific services:**

The required TRUSTS services for UC3 are:

- Actors on-boarding and maintenance
- Metadata catalogue for data and services
- On-boarding of data and maintenance
- Services on-boarding and maintenance
- Monitoring of service performance/performance metrics
- Service usage analysis and billing (service inclusion in the marketplace)
- Data Quality/Data Enrichment/Data Cleansing/Anonymization verification
- Feedback mechanism for users for usage/service delivery, quality, SLA performance, enhancements etc.

The services listed above will be implemented on and will benefit from TRUSTS federated infrastructure, by the following ways:

- Multiple data providers will offer their data, so, especially for UC3 this will be very important for the quality of the results (the more data, the better AI built models).
- Change of business model that will transform an on premise based service to the model of SaaS.
- Standardization of data and their supporting metadata.
- Common services re-use (anonymization, secure data exchange, data correlation, data cleansing/validation/enhancements/standardization etc.).

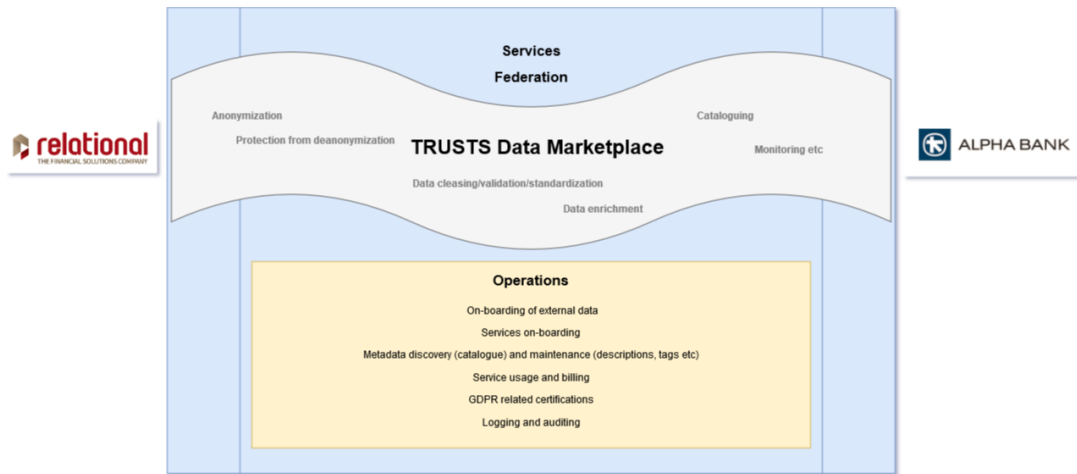


Figure 31: UC3 Structure

### UC3 high level trial scenarios:

Table 9: UC3 high level trial scenarios

UC3 Trial Scenarios	Actors	Evaluation
Actors onboarding	REL	Actor enrolment and verifications for these organizations; actor types: data owners/providers etc; user access management
Actors maintenance	REL	Actors maintenance according to their organisation rights
On-boarding of data	Creditor/End User REL	Data sets on-boarding; privacy options: public/private (restricted) etc; auditing and logging
Data maintenance	Creditor/End User REL	Load new versions of data, incremental loads, deletes etc, structure changes
Services on-boarding	REL	Service on-boarding; terms and conditions associated to the service; contract management
Services maintenance	REL	Load new versions for services; disable older versions; new/disable service

		announcements etc
<b>Catalogue search for data and services</b>	Creditor/End User REL	Discovery of required data and/or services throughout all the federated nodes
<b>Download/Consume data</b>	Creditor/End User	Access data according to the respective smart contract provisions
<b>Consume services</b>	Creditor/End User	Access services according to the respective smart contract provisions
<b>Service usage analysis and billing (service inclusion in the marketplace)</b>	Creditor/End User REL	Evaluate service usage analysis tools; billing and contract; SLA; logging
<b>Security Auditing</b>	Creditor/End User REL	View audit trails of user's activity and data and services' access.

### 8.3.2 UC3 roles

The actors involved in Use Case 3 are:

**Data Provider/End User (Banking Organization/Creditor: Alpha Bank):** Provider of financial/personal data, purchase of anonymised telecommunication customer data and targeted marketing analysis. Anonymization of data for privacy preservation. Data cleaning and pruning to reduce noise and useless entries. Model training and iteration of extraction/cleaning process if needed.

**Service Provider (REL):** Extraction of key data from core banking systems, REL main contribution is to provide and advance in their products. Relational Romania will bring the AroTRON Collection & Recoveries to test and validate improved and more natural ways of communications and debt collection for banks. REL will be the leading partner of WP2 on Natural Interaction and will coordinate the work under this demonstrator. REL is an experienced partner both in terms of coordination and management of collaborative European projects as well as software provider to the finance and banking sector.

**FNET (Tester):** will test the service e.g. providing communication channels from web customers that will allow agents to handle many conversations with end-customers at the same time.

**FORTH (Developer):** Will contribute to the improvement of the conversational UI's usability by conducting evaluation sessions with UX experts. FORTH will also provide anonymization services for the data to be provided by FNET.

To this end, FNET will use the REL service in order to evaluate it for technical usability.

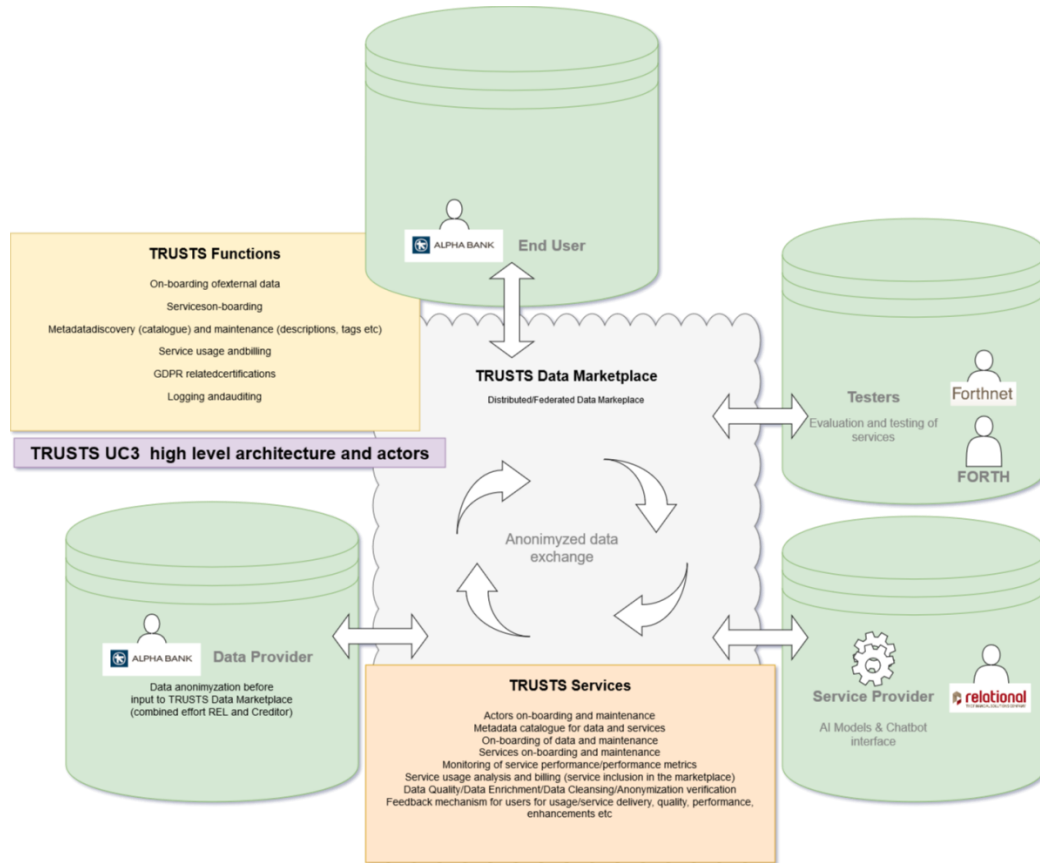


Figure 32: UC3 Actors

The figure above is an illustration of the overall architecture for TRUSTS Use Case 3 that includes all actors involved in Use Case 3 as well as the execution flow diagram.

### 8.3.3 UC3 KPIs

UC3 KPIs for the envisaged data marketplace processes are presented below:

Table 10: UC3 performance and process KPIs

Process	UC3 evaluation KPIs
<b>Actors on-boarding</b>	Alpha Bank, REL, FNET and FORTH actors are successfully on-boarded. Roles are created successfully. Verification of contract fulfilment is done properly.

<b>Actors maintenance</b>	Tests: add representative/disable representative; add new property/update property etc. Example: disable Alpha Romania, add Alpha Cyprus etc.
<b>On-boarding of data</b>	Data sets are on-boarded and configured. Example: Alpha Bank uploads data: financial data sets etc.
<b>Data maintenance</b>	Example: Alpha Bank updates some data sets; deletes other data sets etc.
<b>Services on-boarding</b>	Services are successfully on-boarded and configured. REL is able to provide the 2 services for UC3. (AI Models and Chabot).
<b>Services maintenance</b>	Example: REL adds new service component; uploads new version of services etc.
<b>Catalogue search for data and services</b>	Example: Alpha Bank, end user, searches for Automated Debt Collection service to acquire and for Chabot; return REL provided services. Response time: less than 1s User task success > 90% User satisfaction SUS score>90
<b>Download/Consume data</b>	Example: REL downloads data for the development of the models.
<b>Consume services</b>	Example: Alpha Bank, end user, receives AI and analysis results. Customer loyalty NPS > 8 [0-10] User satisfaction, SUS score > 70 Detailed results analysis, SUS score > 85 Service excellence, SUS score > 80
<b>Service usage analysis and billing (service inclusion in the marketplace)</b>	Example: Alpha Bank, REL can get reports/statics, service usage analysis from the data marketplace.

### 8.3.4 UC3 longer terms business KPIs

The KPIs defined in DoA and the process to meet them is outlined below:

Table 11: UC3 longer term business KPIs

KPI	Baseline Value	Target Value (M36)	Process to meet the target KPIs
Decrease (X%) operational cost for the same collectability	Decrease (estimated at 5%) operational cost for the same collectability	Decrease (estimated at 20-25%) operational cost for the same collectability	A final measurement of KPI needs the solution to be installed at production and run for a period in order to fine-tune and afterwards measure the KPIs.
Increase (X%) efficiency and productivity	The overall contact center efficiency will be increased by 5% with the help of the Virtual Assistant.	The overall contact center efficiency will be increased 15% with the help of the Virtual Assistant.	Final measurement of KPI needs the solution to be installed at production and run for a period in order to fine-tune and afterwards measure the KPIs.
Cost reduction (X%) for process costs on debt management services	Decrease in debt management operational costs (through a 20% increase in process automation).	Decrease in debt management operational costs (through a 40% increase in process automation).	Final measurement of KPI needs the solution to be installed at production and run for a period in order to fine-tune and afterwards measure the KPIs.
Complaints Rate KPI	Decrease of 5 to 10%	Decrease of 5 to 10%	Final measurement of KPI needs the solution to be installed at production and run for a period in order to fine-tune and afterwards measure the KPIs.
Process automation increased (X%)	Estimated increase in efficiency and productivity by over 15%	Estimated increase in efficiency and productivity by over 25%	Base line will be taken during analysis phase from the Creditor, to register current KPI metrics (AS IS) and to be able to compare with new results (TO BE). Final measurement of KPI needs the solution to be installed at production and run for a period in order to fine-tune and

			afterwards measure the KPIs.
Increase (X%) collectability of debt	Estimated increase in collectability of debt by 10%	Estimated increase in collectability of debt by 20%	Final measurement of KPI needs the solution to be installed at production and run for a period in order to fine-tune and afterwards measure the KPIs.
Improve (X%) default event predictability	Foreseeing the end-customer's probability to default in at least 20% of the cases.	Foreseeing the end-customer's probability to default in at least 60% of the cases.	Final measurement of KPI needs the solution to be installed at production and run for a period in order to fine-tune and afterwards measure the KPIs.
Number of data providers interacting with the platform	1 at the start of the use case	Minimum 3 by M36	In order to achieve this the project needs to involve additional data providers using dissemination activities
Number of end-users interacting with the Platform	1 at the start of the use case	Acquisition 3 customers by M36	In order to achieve this the project needs to involve additional data providers using dissemination activities

## 8.4 UC requirements

The consolidated UC requirements for the TRUSTS platform implementation and operation are presented hereinafter.

Please note that the tables below constitute the superset of the requirements from all UCs towards the respective implementation work packages.

### UCR1 Service onboarding environment and procedures are required

#### Remarks

It is required that the TRUSTS data marketplace will have the ability to:

- On boarding of applications in a way that the respective services will be operated within the data marketplace environment
- Liaise through standard API with external applications

Procedures and tools for on-boarding, testing, contracting and becoming operational is required.

#### **UCR2 Data on-boarding and updating environment and procedures are required**

##### **Remarks**

It is required that the TRUSTS data marketplace will have the ability to on-board and update data/metadata. Procedures and tools for data/metadata on-boarding, contracting and becoming operational are required.

#### **UCR3 Unified rich data and service catalogue search transparency for all the federated nodes is required**

##### **Remarks**

The data marketplace user should be able to use a rich search engine to trace the required data/metadata and services querying using a variety of attributes in a transparent way throughout all federated nodes.

#### **UCR4 Procedures of clientele and users on-boarding are required.**

##### **Remarks**

The data marketplace should provide tools for clients on-boarding and contracting. It should provide a wide variety of subscription plans and ad-hoc services. The respective contracting and billing services should be implemented as well. Each client should be able to define users and roles according to the rights provided by the chosen subscription plan.

#### **UCR5 Contract fulfilment, quality monitoring, logging and compliance monitoring is required.**

##### **Remarks**

The data marketplace should implement the tools to verify that contracts have been fulfilled following a transaction of data usage e.g. data provided was of adequate quality according to what had been announced. In addition the users should be surveyed for the perceived quality and for evaluating the data/services that they used. This feedback should be used to improve the data market performance. Feedback on data/services will be used to evaluate the reputation of the respective providers. Transactions will be logged in order to ensure contract fulfilment and prove evidence in case of disputes.

#### **UCR6 Privacy preservation mechanisms are required**

##### **Remarks**

The data marketplace should implement tools for privacy preservation e.g. de-anonymization risk analysis in



order to safeguard that data are adequately anonymised prior to be used in the TRUSTS data marketplace.

#### UCR7 Advanced AI data analytics tools are required

##### Remarks

The data marketplace should offer AI data analytics tools to be used by third party services or directly on data.

## 9 TRUSTS Requirements Analysis and Categorization

Analysing the above described sources a comprehensive set of important functional requirements have been specified. The following table summarizes these requirements, providing:

- the main description of the actual requirement elicited,
- a unique identifier number for each requirement,
- the pertinent current marketplace, questionnaire, interviews and use cases requirement references and
- the involved tasks for the requirement implementation.

Table 12: TRUSTS Functional Requirement Specifications

Req. ID	Description	Requirement reference	Tasks
<b>Datasets and services onboarding functionality and processes</b>			
FR1	The system should provide standardized API descriptions for enabling providers to onboard their datasets and services	SR2, QR3, QR4, QR19, QR20, IR4, UCR1, UCR2	T3.3
FR2	The system should provide APIs that enable its interoperability/federation with other industrial marketplaces and external sources	QR3, QR4, QR20, QR27, IR4, IR5, IR7, UCR1, UCR2	T3.3
FR3	The system should be able to provide datasets and services descriptions	QR20, IR3, IR5, IR6, UCR1, UCR2	T3.4
FR4	The system should provide reference mechanisms to open data from 3 <sup>rd</sup> sources, so as to make available as an option through its data exploration, profiling and provision mechanisms	SR1, QR26, IR6, IR4, UCR1, UCR2	T3.4
<b>Intelligent data/service exploration and correlation functionality and processes</b>			
FR5	The system should provide rich search mechanisms across all federated nodes for available datasets and services	QR4, QR7, QR8, QR28, QR35, IR5,	T3.3, T3.4,

		IR6, IR7, UCR3	T3.6
FR6	The system should be able to provide datasets and services recommendations to its' users pertaining to their profile and needs	QR7, QR11, IR6, IR7, UCR3	T3.6
FR7	The system should employ matchmaking mechanisms through which hosted datasets are matched with hosted services (e.g., suitable for their analysis) and vice versa.	SR3, QR7, QR16, QR11, QR37, QR39, IR5, IR7, UCR3	T3.6
FR8	The system should identify and match related datasets so as to provide combined and enriched data	SR3, QR7, QR24, QR35, QR37, IR5, IR7, UCR3	T3.6
FR9	The system should be able to improve datasets and services profiles based on extracted information originating from the available data	QR7, QR16, QR37, IR5, IR7, UCR3	T3.6
<b>Purchasing and billing</b>			
FR10	The system should provide smart contract mechanisms as a validation means of sellers/buyers agreements	SR4, QR3, QR4, QR13, QR17, IR1, IR3, UCR4	T3.2
FR11	The system should ensure the integrity and authenticity of the smart contracts signed by its users	SR4, QR17, IR1, IR3, UCR4	T3.2
FR12	The system should provide a human friendly representation of smart contracts (e.g., natural language)	QR4, IR1, IR3, UCR4	T3.2
FR13	Signed smart contracts should be legally valid, enforceable and interpretable	QR3, QR17, IR1, IR3, UCR4	T3.2
FR14	The system should encompass mechanisms for keeping transactions performed ensuring that they cannot be infringed	SR4, IR3, UCR4	T3.2
FR15	The system should provide billing mechanisms for enabling consumers to pay providers according to the agreed smart contract.	SR4, QR9, QR13, QR22, QR33, IR3, UCR4	T3.2
FR16	The system must provide alternative and flexible pricing models taking into consideration the diversity of the available datasets and services	SR4, QR9, QR13, QR25, IR3, UCR4	T3.2
FR17	The system should provide brokerage mechanisms for addressing the offerings and demands of the hosted datasets and services	QR4, QR9, QR13, QR40, IR3, UCR4	T3.6
<b>(Meta-)Data Governance</b>			
FR18	The system should provide explicit metadata information for each	SR3, QR7, QR10,	T3.4

	dataset or service is accommodated in the platform	QR25, QR19, QR29, QR36, IR5, IR7, UCR3	
FR19	The system should incorporate models, ontologies and taxonomies for the classification and semantic representation of the accommodated datasets and services in the platform	SR3, QR7, QR11, QR19, IR5, UCR3	T3.4
FR20	The system should be able to incorporate well established or standardized ontologies from different scientific, industrial and business domains for the description of the semantic representation of the hosted datasets and services	SR3, QR7, QR11, IR5, UCR3	T3.4
FR21	The system should provide mechanisms capable to identify the provenance of the hosted datasets	QR4, QR10, IR1, IR3, IR5, UCR3	T3.4
FR22	The system should provide mechanisms capable to identify the lifecycle of the hosted datasets	QR4, QR10, IR1, IR3, IR5, UCR3	T3.4
FR23	The system should harvest metadata extraction from external datasets	SR1, QR36, IR5, UCR3	T3.4
FR24	The system should be able to provide semantic information even for unstructured datasets	QR7, QR36, IR5, IR7, UCR3	T3.4
FR25	The system should be able to keep continuously updated profiles of the hosted datasets and services based on related interactions performed with the system	QR10, IR3, IR5, IR7, UCR3	T3.6
FR26	Dataset discovery should be based on the FAIR <sup>35</sup> principle	QR5, QR7, UCR3	T3.4
<b>Data as a Service and Subscribers management</b>			
FR27	TRUSTS datasets and services should be provided to the users on demand, regardless of geographic or organizational separation between provider and consumer taking into account all potential territorial legislation/regulatory restrictions.	SR2, QR1, QR4, SR5, UCR5	T3.5
FR28	TRUSTS should be able to be deployed as a federation <sup>36</sup> of distributed, interconnected and interoperable nodes.	UCR3	T3.1, T3.3, T3.5
FR29	The system should enable its users to explore data and services openly, providing public descriptions. However, purchased data and services need to be exchanged point-to-point, between the	SR2, QR4, QR14, UCR5, IR2	T3.5

<sup>35</sup> <https://www.go-fair.org/fair-principles/>

<sup>36</sup> Federated architecture (FA) is a pattern in enterprise architecture that allows interoperability and information sharing between semi-autonomous de-centrally organized lines of business (LOBs), information technology systems and applications.

	seller and the buyer. Users should be rated for their quality of transactions.		
FR30	The system should support mechanisms for users' (producers/consumers) subscription opting different schemes (e.g., annual, monthly, etc.) and authentication	SR4, QR15, QR38, IR3, IR6, UCR5	T3.5
FR31	The system should support corporate accounts that fall under one subscription/enrolment per organization	SR5, QR4, QR15, IR3, IR6, UCR5	T3.5
FR32	The system should enable users to create, read, update, and delete (withdraw or make unavailable) datasets, services and user profile records	QR1, QR3, QR4, IR3, IR10, UCR5	T3.5
FR33	The system should provide validation criteria for the new enrolled users, as well as, reputation schemes with regard to available datasets and services.	SR4, QR30, QR31, QR34, IR3, IR6, UCR5	T3.5
FR34	The system should allow consumers to announce their need for specific datasets / services if there are not any available, already.	QR4, QR23, IR6, UCR5	T3.5
FR35	The system should provide notifications regarding datasets / services updates to users that have granted access to them	QR4, QR23, IR6, IR7, UCR5	T3.5
FR36	The system should provide easy to use UIs (ensuring effectiveness, efficiency and user satisfaction) that will help users to accomplish their tasks effectively and prevent them from committing errors.	QR1, QR4, QR6, QR14, QR19, QR43, IR5, IR10, UCR5	T3.5, T5.2, T5.3
FR37	TRUSTS UIs and workflows have to follow a business-wise rational (e.g., one stop shop), for coherently mapping the market's needs.	SR2, QR1, QR4, SR5, UCR5	T3.5
<b>Data protection</b>			
FR38	The system must provide cryptographic and secure protocols for the analysis of sensitive data as required by the respective stakeholders.	QR2, QR12, QR42, IR1, IR9, UCR6	T4.1
FR39	The system should provide de-anonymization attack assessment and risk analysis for the private / sensitive datasets to be onboard	QR2, QR12, QR41, IR1, IR9, UCR6	T4.3
FR40	The system should employ anonymization tools and guidelines for data anonymization	QR2, QR12, QR32, IR1, IR9, UCR6	T4.3
FR41	The system should provide means for converting algorithms that might compromise the data privacy into safe privacy preserving ones without harming their functionality <sup>37</sup>	QR2, QR12, QR42, IR1, IR9, UCR6	T4.5

<sup>37</sup> Such requirements are set by the relevant stakeholders. Please note that the functional requirement does not impose any specific solution. Respective solutions will be evaluated by the technical workpackages of the project.

Advanced data analysis based on Machine Learning			
FR42	The system should incorporate well established ML algorithms that can be used by the TRUSTS customers for data analysis and classification.	QR21, QR39, IR8, UCR7	T4.2
FR43	The system must incorporate a secure infrastructure for the distributed analysis of data based on ML approaches	QR21, QR39, IR8, UCR7	T4.4
Trusted and legitimate data flows			
FR44	<p>Mechanisms provided by the TRUSTS platform regarding personal data, non-personal data and services exploration, exchange agreements and purchase, should be compliant with the following regulations (when applicable):</p> <ul style="list-style-type: none"> <li>• General Data Protection Regulation</li> <li>• e-Privacy regulation, for electronic communications</li> <li>• Free Flow of Non-Personal Data Regulation, for data exchange between the TRUSTS platform and subscribers</li> <li>• Platform-to-Business Regulation, for safeguarding TRUSTS' operational transparency and fairness.</li> </ul> <p>Mechanisms provided ensuring that local laws apply to each federated node.</p> <p>Predefined contracts should exist.</p>	SR4, QR3, QR18, LR1, LR2, LR3, LR4, UCR5	T6.1, T6.2, T6.3, T6.4

## 10 Conclusions and Next Actions

In this deliverable we analysed key sources to gain requirements with respect to the development and operation of an industrial data marketplace targeting the telecommunication and financial sector and beyond, following a comprehensive methodology.

In today's interconnected world business silos seem to fade enabling business expansion to other domains and collaboration of varying industries e.g. use another company data with correlation to own data to achieve better analysis, etc. A forward looking data marketplace endeavour such as TRUSTS aims at raising all obstacles in the data exchange process establishing a comprehensive platform incorporating all respective best practices, standards and regulations in order to address a wide variety of industrial clientele including telecommunication companies, financial institutions and their collaborators.

To this end, an extensive list of Functional Requirements' specifications was produced indicating both the source requirement and the task that will undertake its evaluation and implementation.

Within this deliverable the three Use Cases have been thoroughly analysed in terms of their needs with respect to the TRUSTS data marketplace. Their anticipated operation through the TRUSTS environment is detailed along with the roles, trials descriptions, high level scenarios and respective KPIs.

This set of the information along with the functional requirements will be used for the evaluation of the implementation in order to improve the TRUSTS platform using the task 1.3 methodologies.

This deliverable constitutes the first version of the two reports containing the detailed analysis of the requirements for a commercial financial and operators' industry vertical data marketplace platform and the use cases definition including the target KPIs that would set the benchmarking for the actual measurements.

Key strategic outcome from the analysis of the elicited requirements from all sources is that the overall TRUSTS objectives are in line with all the key stakeholders' expectations thus setting the bar high for defining a successful service having significant impact on the data industry.

It should be noted that due to COVID-19 lockdown and the limited time since the project commencement a few number of responders provided feedback to the questionnaire. Nevertheless, the questionnaire responses along with the interviews provided a solid ground for the identification of the stakeholder requirements. The outlined Functional Requirements are technology agnostic since they do not aim to set the implementation framework but rather the required functionalities and processes of the TRUSTS data marketplace. Task 2.4 will define the architecture principles while the end to end environment will be tested through the UC trials.

Task 2.2 will continuously collaborate with:

- All WP1 tasks in order to evaluate additional information with respect to the TRUSTS architecture and data marketplace initiatives and trends.
- WP7 which will produce adequate business models and receive market feedback from any related exploitation action
- WP6 which will define the legal aspects and processes
- WP3 and WP4 which will undertake the platform development
- WP5 which will execute the UC trials providing valuable feedback and
- WP8 to provide information and evaluate feedback from respective events.

The aim is to systematically assess the input from all designated sources in order to update the requirements driving the development of an imminently exploitable TRUSTS platform.

It is understood that the FRs and UCs are aiming at defining the functionality and operational requirements of the end-to-end platform. Nevertheless, an analysis will be made in the technical WPs in order to evaluate which functionality can be feasibly implemented within the resources of the project. The implementation tasks aim at producing an environment that will be able to support all essential data marketplace functionalities. In any case, it should become clear that some functionality which is requested by the FRs cannot be fully addressed within the scope of the project and according to the provisioned resources allocation; this will be appropriately documented and reported, towards the goal of scheduling the implementation as part of the commercialisation phase.

This work will be thoroughly analysed in the deliverable D2.3 entitled “Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II” which is due on M24.

Towards advancing T2.2 activities and the implementation of the D2.3 deliverable the project will assess the initial developments, identify additional stakeholders and refine questionnaires in order to constitute the final set of the TRUSTS environment requirements.

# Annex I: Questionnaire

## TRUSTS requirements elicitation questionnaire

### 1) Inform & Consent

*You are invited to take part in the TRUSTS project questionnaire. Your participation is voluntary and you may decide to withdraw it at any time.*

*The purpose of the TRUSTS research project is the development and testing of a federated data marketplace. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871481 and will last 3 years until the **31 December 2022**.*

*The TRUSTS consortium aims at receiving responses to the questionnaire and interviewing industrial, academia and regulatory domain experts in order to lead the TRUST data marketplace specification. Your responses will help us to evaluate the functionality, services and operational capacity of such an endeavour and to establish its operation.*

*None of the personal data acquired will be disseminated or distributed outside the TRUSTS consortium. More information with regard to the TRUSTS research policy can be obtained from the project coordinator:*

*Alexandra Garatzogianni - H2020 Coordinator of TRUSTS Trusted Secure Data Sharing Space, Senior Project Manager, Leibniz University of Hannover, L3S Research Center & Head of Tech Transfer, EU-Project Coordination & Management, Leibniz Information Center for Science and Technology, University Library*

*Declaration of consent to participate in the research questionnaire:*

*By agreeing to answer this questionnaire I accept that, I have read and agree with the Consumer pilot participation rules and regulations as they are described below:*

- This survey is being performed as part of a research project.*
- The data I provide will be used only for research purposes.*
- The data I provide may be published internally or externally and be used as part of presentations related to the research. Any publication of the data will be in an anonymised form with all identifying personal information removed.*
- I may withdraw my consent to participate in this research questionnaire at any moment.*
- Personal data provided are limited to the identification of the respondent (name/title) and to their place of employment. Such information are processed based on **the legitimate interest of the TRUSTS consortium, namely to conduct scientific research as described in the document.***

***I understand what the study is about and my questions so far have been answered. I agree to take part in this study\****

☐ Yes

### 2) Name of your organisation?

---

### 3) Please specify your sector

☐ Public

☐ Private: \_\_\_\_\_ \*

☐ Academic



☐ Regulatory body

4) Please indicate the approx. size of your organization (no of employees)

☐ 1-50

☐ 51-250

☐ 251-1000

☐ more than 1001

5) What is the main country of your current workplace?

---

6) Which of the following roles fits you best? (multiple options allowed)

☐ Business driver

☐ Strategical driver

☐ Technical driver

☐ Domain Expert

7) Please indicate your management level

☐ Executive officer (i.e. CEO, CTO, CFO, COO etc.)

☐ Operating officer (i.e. General manager, Plant manager, Regional manager, and Divisional manager etc.)

☐ Administrative officer (i.e. Office manager, Shift supervisor, Department manager, Foreperson, Crew leader, Store manager, Project leader etc.)

☐ Professor

☐ Researcher

☐ Other - Write In: \_\_\_\_\_

8) Please indicate your years of business experience

☐ None

☐ Less than 2 years

☐ 2 to 5 years

- ☐ 5 to 10 years
- ☐ 10 to 20 years
- ☐ 20+ years

9) I understand the data sales/buying process of my organisation (1: totally disagree, 5: totally agree)

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5

10) I am involved in buying and selling of data in my organisation (1: not at all, 5: fully involved)

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5

11) Please describe the role of your organisation (multiple options allowed)

- ☐ Data buyer
- ☐ Data provider
- ☐ Application/service provider
- ☐ Application /service user
- ☐ Data marketplace platform operator
- ☐ Standardisation body/regulator
- ☐ Other - Write In: \_\_\_\_\_

12) Please tell us how often your organisation needs to buy data (per month)

13) Please describe the type of data your organisation needs to buy

---

---

---

---

14) Please describe the desired process to buy data

---

---

---

---

15) Please tell us how often does your organisation needs to provide/sell data (per month)

---

16) Please describe the type of data that your organisation needs to provide/sell

---

---

---

---

17) Please describe the process you would like to follow to provide/sell data

---

---

---

---

18) Please tell us how often does your organisation provides data applications/services (per month)

---

19) Please describe the applications/services that your organisation provides

---

---

---

---

20) Please describe the process for providing applications/services you would like to follow

---

---

---

---

21) Please tell us how often your organisation buys data applications/services (per month)

---

22) Please describe the applications/services you buy

---

---

---

---

23) Please describe the process for buying applications/services you would like to follow

---

---

---

---

24) Please tell us how often your organisation provides data services through your platform (per month)

---

25) Please describe the services that your platform provides

---

---

---

---

26) Please describe the process for providing services through your platform that you would like to follow

---

---

---

---

27) Please describe the standardisation status with respect to data exchange between different organisation and data marketplaces

---

---

---

---

28) Please describe in your opinion the standardisation gaps and the way forward to boost the data marketplace endeavour/Please describe the required standardisation for federated data marketplaces

---

---

---

---

29) Data marketplaces can help making the process of selling and buying data easier. Which services or functions would you need from a data marketplace? (multiple options allowed)

- ☐ Anonymization
- ☐ De-anonymization
- ☐ De-anonymization risk analysis
- ☐ Data hosting
- ☐ Metadata hosting
- ☐ Datasets catalogue
- ☐ Datasets valuation
- ☐ Datasets rating

- ☐ Third party Applications/Services catalogue
- ☐ Datasets search discovery service based on metadata, ontologies, etc.
- ☐ Datasets/Applications trading service
- ☐ Multiparty computation (different parties to jointly compute a function over their inputs while keeping those inputs private)
- ☐ Private set intersection (a cryptographic technique that allows two parties holding sets to compare encrypted versions of these sets in order to compute the intersection)
- ☐ User authentication
- ☐ User role rights according to GDPR processes
- ☐ Transaction logs
- ☐ Billing
- ☐ Standardisation information
- ☐ Other - Write In: \_\_\_\_\_

30) Please indicate which of the functionality you selected in the previous question is mandatory for a data marketplace (multiple options allowed)

- ☐ Anonymization
- ☐ De-anonymization
- ☐ De-anonymization risk analysis
- ☐ Data hosting
- ☐ Metadata hosting
- ☐ Datasets valuation
- ☐ Datasets rating
- ☐ Third party Applications/Service catalogue
- ☐ Datasets/Application trading service
- ☐ Multiparty computation (different parties to jointly compute a function over their inputs while keeping those inputs private)
- ☐ Private set intersection (a cryptographic technique that allows two parties holding sets to compare encrypted versions of these sets in order to compute the intersection)
- ☐ User authentication
- ☐ User role rights according to GDPR processes
- ☐ Transaction logs
- ☐ Billing
- ☐ Standardisation information
- ☐ Other - Write In: \_\_\_\_\_

31) Please describe the advantages of a data marketplace

---

---

---

---

32) Please describe what a data marketplace should avoid

---

---

---

---

33) What pricing model for a data marketplace would you prefer? (multiple options allowed)

☐ Free without service level agreement

☐ Fixed price subscription

☐ Package

☐ Pay per use

☐ Progressive price

☐ Other - Write In: \_\_\_\_\_

34) Optional: Would you like to tell us anything else in relation to the data marketplaces?

---

---

---

---

35) Optional: Please describe a data marketplace use case?

---

---

---

---

36) Would you like us to contact you to further discuss your view on the data marketplaces

( ) Yes

( ) No

37) Name:

---

38) Title:

---

39) Organisation:

---

40) Please provide contact details (tel/email/skype)

---

---

---

---

41) Preferred interview method (tel., skype, face2face, etc.)

---

---

Thank You!



## Annex II: Interviews guidelines

The interview should be structured around the same questions we had in the questionnaire but we have to have an open discussion trying to deep dive into the interviewee answers by requesting clarifications or examples when he/she responses for a company process, data exchange, applications needed, etc.

We do not to focus only to GDPR, privacy, authentication, security or even anonymized data since we believe they will all refer to them.

Assuming that all the above are in place, what could make organizations use a data marketplace e.g. do they find datasets in the business neighbourhood or they have to search far way and why e.g. compare patterns, etc., how do they value the datasets quality prior to procuring them, do they facilitate their production value chain through datasets search e.g. search for the right material through descriptors, etc.

At the end of the day, we would like to find out the reasons that could drive organizations and employees within these organizations, making a data marketplace part of their everyday business like being yet another utility.

This will provide us enough information to produce and propose to the market a sustainable service (We prefer not to use the word platform because it mainly has a technical point of view. Rather, we prefer using the word service which describes the value that the client receives.).

The proposed duration of each interview is approximately 30 minutes.

## **Annex III: Interviews**

In the following the interviews' summaries and related information reside:

## Interview 1

**Means:** Skype

**Interviewee affiliation:**

Forthnet/Marketing/Online PayTV services Data analyst, More than ten years for experience

**Key outcome:**

The main endeavor is to design new online PayTV products.

In addition, the employee helps designing the specifications of the data that are collected through the set top box (STB) of the FNET clients that are used to access FNET PayTV services from android, IOS and Windows devices. Recently through Android TVs as well.

Current status: Data from online services are recently been collected and are analyzed by own means.

The needs from a data marketplace are the following:

Find similar datasets from external operator

Have access to pattern identification applications based on AI

Have access and possibility to test/simulate recommendation engines

Have the possibility to ask for bespoke data analysis services development and run respective tenders through the data marketplace.

Have vertical sub data marketplaces e.g. for content creation

Accommodate film making value chain<sup>38</sup> with IPR protection:

- Data exchange and warehousing
- Shared data repositories
- Analytics and marketing applications
- Etc.

In addition, there is a potential trend for ontology based recommender systems<sup>39</sup>.

Transaction and participants should be rated.

Discussion forums should exist.

The envisaged interaction frequency is twice a month or more.

Services provision should be based on company subscription (with a certain number of access licenses) and the possibility to procure add-on premium valued added services.

---

<sup>38</sup> A CREATIVE WORKS ONTOLOGY FOR THE FILM AND TELEVISION INDUSTRY (<https://movielabs.com/wp-content/uploads/2018/09/A-Creative-Works-Ontology-for-the-Film-and-Television-Industry-Final-2018-9-24.pdf>)

<sup>39</sup> A Novel Ontology-based Recommender System for Online Forums ([https://www.researchgate.net/publication/236001999\\_A\\_Novel\\_Ontology-based\\_Recommender\\_System\\_for\\_Online\\_Forums](https://www.researchgate.net/publication/236001999_A_Novel_Ontology-based_Recommender_System_for_Online_Forums))

## Interview 2

**Means:** Skype

**Interviewee affiliation:**

Forthnet/Marketing-Strategy, 20 years of related experience in the same position in FNET

**Key outcome:**

Everyday work is mainly focused on:

- Actual vs Budget KPIs tracking for all services of all services (more than 20), all markets (retail, SME, Business/Corporate, Public).
- Benchmarking vs international trends (e.g. bundling services uptake projection, call minutes consumption, viewership trends, bandwidth consumption trends, opex trends, capex).
- Projection of uptake in various markets e.g. MVNO, OTT, Android TV, etc.
- Respective reports production to CEO.

Data used:

- Internal historical aggregated and anonymized data from the marketing department, technical department, HR, finance.
- Projection analysis from respective consultants (e.g. Mason Analysys, IDC, etc.)
- Bespoke projection analysis and consultancy services.

Need:

- Have easy and updated information about new trends data per market and region.
- Have access to sample datasets prior to procuring
- The data marketplace should announce each time a new dataset is uploaded according to the user's profile
- The user should be able to announce to the marketplace needs for datasets. This should be done in a systematic format (not free text) in order to each search e.g. keywords, timeframe of the required data projection, etc.

Data marketplace service model:

- The envisaged frequency of having such services is once every two months.
- A basic subscription model is envisaged with add on packages according to the datasets that are acquired.

### Interview 3

**Means:** Skype

**Interviewee affiliation:**

Forthnet/Marketing, Customer base management (More than five years of related experience, 2 years in FNET.

**Key outcome:**

Everyday work is mainly focused on:

- Actual subscribers base per service vs budget (subscribers no per service, migration between services, updates, ARPU, CLV, churn, retention campaigns, Call center interactions, etc.)

Data used:

- Access to new tools and methods for data analysis e.g. based on AI.
- Trend analytics/dataset on similar markets
- Possibility to predict behaviour using historical data (e.g. downgrades if a new pricing policy is enforced).

Need:

- Easy access to new algorithms and applications as opex and not as capex.
- Similar markets datasets.
- Correlation of data with other related industries could be considered positively if done in a standard manner and easily.

Data marketplace service model:

- The envisaged frequency of having such services is multiples time each month.
- A basic subscription model is envisaged.

#### **Interview 4**

**Means:** Skype

**Interviewee affiliation:**

Forthnet/Legal Department

**Key outcome:**

**A.**

- The data marketplaces as all electronic transactional means should promote trust.
- Proposal to have clear terms of use visible to the users fall all actors.
- Visible rule for personal and non-personal data exchange according to the regulations
- Investigate possible legal differences in the federated nodes.
- Feedback reception, analysis and response mechanism should be established for all processes
- Mechanism to evaluate and publish the reputation of each actor that performs a transaction e.g. evaluating previous transaction and quality of current offered datasets or applications.

**B.**

- A set of contracts for data usage, services, revenues split etc. should be established with specific terms. Avoid negotiating every contract separately.

## Interview 5

### IDSA

The IDSA cannot be called an operator of a data marketplace because we neither own nor run one. We are rather working on the development of a Reference Architecture as a model for any ecosystem where data can be exchanged between two or more parties. Therefore, we can answer your questions only on a theoretical level:

- Description of current process of onboarding applications/services and needs for improvement (e.g. application compliance rules/standards, testing process/period, SLA, contract negotiation between the application/service owner and the data marketplace, smart contracts, life cycle, T&Cs for using the application, data type used, etc.)
  - We have no practical experience in such processes because we do not run a data marketplace. We can only say from a theoretical perspective, that we see a need for a certification process for all organizations and also all their components that will be plugged into an IDS ecosystem. The certification process is theoretically written down in this document, but have not been finished or proven yet. In order to gain the digital certificate x.509 the participant and its component will have to go through the process described in the previous document. The IDSA is also working on a pre-test, if the company and its components are compliant to the requirements by “visiting” a test facility, where they can plug their components to a test infrastructure. However, currently we are setting up such a test facility with a company called SQS.
  - Further, we advocate the concept of usage control IDS Reference Architecture Model 3.0 (from page 83), which is an extension of data access control and enables security requirements that cannot be achieved by access control only (secrecy, integrity, time to live, anonymization by data aggregation, anonymization by data substitution, separation of duty, usage scope).
- Description of current process of onboarding data/metadata and needs for improvement (e.g. SLA, contract negotiation between the data/meta data owner and the data marketplace, smart contracts, life cycle, T&Cs for using the data, descriptors, standards, etc.)
  - We have not defined the negotiation of contracts since we are not a data marketplace.
- Data marketplace service/data catalogue and search facilities (introduction of services and data to the catalogue, catalogue publication process, search functionality, etc.)
  - We have developed the theoretical models of a data broker and an appstore which are written down in the IDS Reference Architecture Model 3.0 (from page 67).
- Data marketplace compensation process (e.g. how the fee for using a service/dataset is allocated to the service/dataset owner and the marketplace operator, etc.)
  - We do not have an approach for this yet.
- User/Subscriber enrollment process (roles, etc.)
  - We do not have an approach for this yet.
- Marketing activities/expected market growth/niche vs mass market
  - Since we are not an operator this is not applicable for us.

- Roles in operations
- Infrastructure/architecture/availability/SLA maintenance
  - The IDS Reference Architecture Model 3.0
- Other
  - From our overriding perspective we may say that the demand for another data marketplace is not strong enough. We also observed that there is a need for overall data infrastructure for essential and basic services (digital identity and trust management) and a governance to create and maintain trust between all involved participants.



## Interview 6

Means: Skype

Interviewee affiliation: Corporate – Legal Advisory Services, & Financial Compliance and Reporting Services. More than 15 years of experience.

Outcome:

Everyday work mostly focuses on:

- Provision of Corporate Services to their clients (more than 20 clients) in several market areas
- Provision of Financial & Compliance Reporting Services
- Facilitation of production value chain through datasets search
- All services provisions are based on client's request

Current Status:

- Data collection in a weekly basis, analysed by company's private means.
- Purchase & Sale of data in a weekly basis
- Purchase of data from open corporate services upon request of their clients. Request to the open corporate provider, checks for any suspicious behaviour and alerts for any malicious behaviour of client's data and analysed data ends up to the end-user.

The needs for a data marketplace are as mentioned below:

- Comparison of patterns in an advanced data marketplace
- Securely and GDPR compliant exchanged market data to the clients
- Data readiness for correlation
- Fast and comprehensive use access to data market applications
- Better suspicious activity report structure
- Prospect to buy up add-on premium valued added services.

## Interview 7

Means: Skype

Interviewee affiliation: Audit, Taxation & Business Advisory Services. More than 25 years of experience.

Outcome:

Everyday work mostly focuses on:

- Provision of financial services for their clients (more than 30 clients) in various markets such as corporate, banking, legal etc
- Provision of consulting services to their clients
- Services provision situated on client's requests

Main issue is to provide securely market data to their clients and GDPR Compliant

Current Status:

- Data from open corporate services providers are collected regularly and are analysed by company's own tools.
- Purchase & Sale of data in a monthly basis
- Purchase of data from open corporate services, via an annually subscription and not from a data marketplace that data will be combined and securely checked

The needs for a data marketplace are as mentioned below:

- Easy and friendly use access to applications related to data market
- Number of data providers interacting together for better results
- Securely exchange of data related to market data
- Possibility for addressed data analysis services development and run various tenders via the data marketplace.
- Better detection accuracy of data

## Interview 8

Means: In person

Interviewee affiliation: IT Development & Business Corporate Services. More than 15 years of experience.

Outcome:

Everyday work mostly focuses on:

- Provision of technologically-advanced e-business software solutions (more than 30 clients) in various markets such as corporate, banking, legal, audit etc
- Provision of a state-of-the-art comprehensive software suite which manages corporate administration, risk analysis, compliance, tax reporting, and much more.
- Services provisions are based on client's requests
- Services provision of maintaining and fulfil regulatory compliance mainly for AML activities and purposes
- Provision of Transaction Monitoring services so as to minimize the risk of financial fraud and terrorism financial activities.

Current Status:

- Purchase & Sale of data in a daily basis
- Possibility to predict behaviour using datasets.
- Daily collection of data from open corporate services providers and analysation by company's means.
- Upon client's request, access to the corporate provider, checks for unnormal behaviour and malicious patterns, alerts for any suspicious behaviour of client's data and analysation of results ends up to the end-user

The needs for a data marketplace are as mentioned below:

- Artificial Intelligence and Machine Learning techniques for better detection and accuracy determining the risk level and mitigating the several exposures and associated with the onboarding and monitoring of the customers
- Dynamic data marketplace with advanced rules-based engines for fast and exactness investigations for suspicious cases
- Customers confidence and satisfaction by offering high quality data services
- Increase of data providers interacting together for better results
- Increased reliability
- Raise productivity and enriched cost-effective functionality
- Decrease ownership and operational costs
- GDPR aligned and saved exchange of data related to market data

## Interview 9

**Means:** Microsoft TEAMS

**Interviewee affiliation:**

Alpha Supporting Services.

Project leader to a company provides IT services in the fields of banking applications (e.g. core banking, loan collections, AML)

**Key outcome:**

The main scope is to support the advancement of the data economy support and benchmarking of collections operations. Furthermore, using the exchange of data, through the Data marketplace, the bank will be able to follow the technological evolutions for supporting the processing of huge amounts of complex and diverse data in real-time. By providing its data sets that are growing so large and complex, to Fintech SME's it will enable the development of models and mechanisms that will overcome the traditional tools are no longer able to process this data at sufficiently low cost and in reasonable time. Finally, in this way we can help financial services providers grow, innovate and rapidly deliver Technology and customer value.

Recent success of the Fintech robo-advisors, offering automated digital investment advice using their gathered customer profile information, shows that Fintechs are already able to convert Big Data into new compelling customer services.

Current status: The Company supports Big Data for the segmentation of customers, based on the available data (e.g. customer profiling, analysing transaction patterns, past and immediate customer behaviour) to get real-time customer insights.

The needs from a data marketplace are the following:

- Under a marketplace approach, the bank can provide its data up to anyone who is interested, including other Fintechs, and third-party developers, so that they can utilise the bank's data to build their own products to be accessed via a platform.
- A vast quantity of data (i.e. terabytes or petabytes) to be handled. These huge amounts of data make it impossible to be processed by traditional data processing tools within reasonable time delays.
- Big Data technologies should be able to process both batch and real-time data. For real-time data, quick analysis for (near) real-time insight generation can be a necessity for the business.
- Multiple types of data should be supported, i.e. from highly structured data to unstructured info like text, video, audio, blogs, tweets, Facebook status updates
- Through the exchange of data in the marketplace it will be allowed also to **deliver new, innovative products and services to customers**, which use the insights derived from the data streams.

The Data Marketplace should also offer modelling and analysis using a variety of techniques to model and analyse data and often techniques are combined to get the best result. In analysis the following groups would be useful to be identified

- Unstructured data to structured data conversion: these techniques transform unstructured data to structured data, on which other Big Data techniques can be applied. This includes "Text Analytics", "Picture Analytics", Audio Analytics" and "Video Analytics".

- (Descriptive) Data Mining: this refers to techniques using algorithms to discover hidden patterns, relationships, dependencies and unusual records or dependencies.
- Predictive analytics: a variety of techniques to make predictions (determine likelihood of future events, i.e. future trends or likely behaviour) from historical and current data patterns. Often based on time-series analysis, this type of analysis is typically used for determining the "next-best-offer" and implementing adaptive user interfaces.
- Machine learning: this group of techniques consists of applying one of the above techniques but adding the element of automated learning to it. This means the analytic technique will learn itself to provide better insights into the data, i.e. the model compares expected outcome with the real outcome and adapts accordingly to better align for future predictions.

In addition, there is a potential need for:

- Anonymization
- De-anonymization
- User authentication
- User role rights according to GDPR processes
- Transaction logs
- Datasets catalogue
- Billing
- Private set intersection (a cryptographic technique that allows two parties holding sets to compare encrypted versions of these sets in order to compute the intersection)

## Interview 10

**REL**

**Location:** Skype

**Debt Collection Law Firm (Debt Servicer), interviewee:** Operations Manager, with more than 10 years of experience in the field

**Daily responsibilities:** The Operations Manager overviews the business lifecycle by analysing performance data and takes decisions on possible improvements and opportunities or new areas of business process improvements.

**Business Description:**

*Debt collection law office primarily collect debts by making coordinated contacts with debtors in order to formulate a commonly agreed plan for the debtor to pay off all or a portion of the debt owed. The formulation of the space of possible solutions that are both satisfactory for the Creditor and the Debtor is one of the greatest challenges facing my position.*

*The company also employs pre-legal and legal consultants (lawyers), in order to follow a litigation process in the areas where the debtor is not co-operating for an amicable solution. Legal departments handle the needed paperwork and provide required legal services so that the debtor knows that a court date on the matter is pending. The debt collection office firm can take additional action once it has a judgment, including garnishing the debtor's wages, placing a lien on un-exempted property and collecting profits from rental or business income. The challenges of the litigation process is to decide which cases should proceed since this path involves signification overhead costs that in a lot of cases are covered by the Creditor when the Debtor cannot pay them.*

### Interview results (needs and benefits)

There is a strong need to have benchmark data, **representative** and **unbiased** data on debt collection KPIs. Currently KPIs are provided by the Creditors but cannot be either representative or unbiased.

There is a big Interest in **skip tracing data**. Skip tracing is a data-based process to search and locate new contact information for a person/company; currently, this process is very challenging for the company due to the GDPR policy enforcements. GDPR-compliant skip-tracing data would strengthen the debt recovery process by reducing costs, automating more tasks and optimizing the resources. Information such as a debtor's workplace, home place or business name and location is needed. With access to better consumer information, the company can also gain useful insights such as known relatives and associates; and new address details.

There is a strong need to have **standardized** List-Of-Values/dictionaries/metadata/processes, like standard categories and metrics, for example, for the debt collection process. This is very important for best results and comparing the results with industry standards.

There is a need to have **standard entities (metadata and ontologies)** at the level of data marketplace – data governance. **Need for data governance and metadata management**, in general.

All this data management must be **GDPR compliant**.

There is a need to have access to **Credit Bureau/Credit Scoring data**, for example, for both individuals and companies, for financial data; financial results to be joined with other data sets, for more detailed and complete debt information.

Also, there is a need to have **accurate Geographical Information System (GIS) and regional data and mapping of this data to authorities** like: land registries, cadastral institutions, courts etc.

Need to have, also, notifications on data changes; for example, when there is a change for a court (merges between different courts etc.) and how the new mapping will look like.

There is a need for **behavioural history for customers**, in terms of loan lifecycle, where the list of activities performed by both-sides, Debtor and Creditor are listed and codified and also linked with the related financial impact that these actions had (e.g. send a letter → payment by the Debtor).

There is a need for data refresh:

- Benchmark Data should be refreshed monthly since they have important temporal deviations through the year;
- Skip Tracing Data should be accessed daily or even multiple times per day and also to be available on demand;
- Credit Bureau Data should be accessed on a daily basis for both individuals and companies; dictionary metadata are slowly changing (rarely, few times a year).

## Interview 11

**Interviewee: City of Vienna**

**GoToMeeting; Friday, April 3, 2020**

**Additional questions:**

- 1) How many providers do you contact on a regular basis? Are they national and/ or international data providers?**

Interviewee: We do not contact a lot of providers, because as a local government, we have our own data, we have stamp data, we have metrological data from national data providers, but we do not buy a lot of data

- 2) Do you have a need to buy more data?**

Interviewee: Yes, we are speaking about data from business, especially mobile phone data from our national providers, google data. We are working on a strategy belonging this because we think we will not buy this data, we eager think of a data room where all the providers could work together.

- 3) Can you tell us more about the price of the data?**

Interviewee: I do not know the price of the data of weather...I think we have 10 different contracts from 60 Euro per year to I do not know how much per year. But I think it depends on the data.

- 4) Can you please describe in more detail your current process of acquiring data?**

Interviewee: It is not an easy process. If you have a requirement of data, we make a contact to the provider, we have to make a contract with the provider and then we get the data either via E-Mail or API, if possible. I think it is a long way...Time from the first contact to receive the data: We are now in a process from the national metrological institute, which currently already takes 4 months.

- 5) What obstacles and improvements can you see within this process?**

Interviewee I would like to use the Data Market Austria for it. – But I do not know if I can get the data in that way.

**So you expect from a data market to make this process much faster?**

Yes, yes, sure, the process should be faster and easier: from one day to another. I think the meta data – I want to see the price, I want to make the contract online and the next day, if the contract is fixed, I get the data.

- 6) How would look like the ideal search process to working a data market?**

As I said, if I find the data which has a good description via the meta-data, there is a process, like a shopping card, where I can put it in, like on Amazon and any other shopping provider, like this I can shop the data.

- 7) Can you specify the approx. share of each different kind of data (geo, points of interest, statistical and real-time data) you provide / sell?**

Interviewee: We provide the data. We provide the points of interest, but we provide it via open data.

**Which data provide you the most and which data do you provide the least?**

Interviewee: We have 500 different data sets, so we provide 500 different ones...

- 8) What expectations do you have from a data marketplace?**

Interviewee: To have this Data broker functionality. Real good Metadata and maybe those APIs to get the data faster.

- 9) Which services or functions would you need from a data marketplace?**

- 10) What do you expect from, easier access to data thanks to a data marketplace?**



Interviewee: As said, to get the data faster and to know what data is available.

**Please specify what pricing model(s) you chose and why it is your preference.**

Interviewee: From the side of the government it would be easier to have fixed prices. We have a yearly budget and it would be easier to plan if we had fixed prices.

**11) Can you please elaborate in more detail your optional responses when it comes to relation to data marketplaces?**

→ Response in the questionnaire (Standardization, only one per country, harvesting of metadata)

Interviewee: Standardisation would be found if we had kind of a meta-data standard, maybe we could a European standard or the standards which are available. I would prefer to have the Data Market Austria and then harvesting a European Data Market Portal.

**Why would you buy from the Data Market Austria and not from another market in another country?**

Interviewee: I think we have use from the side of the government, we use Austrian data.

**So actually if you need data from another country, you would buy them?**

I do not know what data we would need it from another country. It is not our use case. But maybe it is good to have a European portal but I know it is not easy. It is not easy to have all stakeholders in Austria at one table and to make this collaboration. I would make step by step. If we have a Data Market Austria, which is working, the next step would be to get together with other national markets.

**But when it comes to the data sets you provide? Are they used from international entities?**

Interviewee: Yes, they are used internationally, for example from Russia or in the US, I think. Especially the points of interest, the real-time data is used from visits from China, Russia, the US or somewhere else.

**12) (Can you share your knowledge about Data Market Austria and the potential involvement/ expectation of your entity?)**

Interviewee: My expectation is to find data easier at the moment. I know the platform but I cannot find data here.

**Questions from the interviewee:**

- Are there companies interested in it? Three use cases.
- Asks for telecommunication companies? Yes, one but not mobile options.
  - Interesting because there are lot of entities interested in selling the data.
- Question about the timeline of the project.

## Interview 12

**Means:** Skype

**Interviewee affiliation:**

Bank/Digital Business/Customer Innovation – Senior Manager

**Key outcome:**

Digital Innovation is an important aspect of Digital Business in the Bank. We are responsible for implementing new services for internet & mobile banking, new machines, personal financial management and digital innovative services. Internally we use data analytics in order to design new ideas & business. We also use our Customer Value Management department in order to make data mining & complex data sharing. We think it would be important for the bank to use applications for machine learning.

A major current task is the deployment of the e-branch stores of the Bank, which aim to shift the current transaction process of the Bank's customers towards e-services under the overall concept of moving the services to their digital transformation.

Although the banking system, due to its nature, is quite introvert with regard to getting into transactions with external data and services, recently some steps towards cooperating with 3<sup>rd</sup> party organizations have been taken, so that new added value products that combine services from both the bank and the organizations are created.

To that end, the Bank has recently started to pursue digital markets in order (a) to reinforce and improve its internal digital services (such as the financial analysis and risk assessment) and (b) to identify potential cooperation with 3<sup>rd</sup> party organizations.

Thus, according to recent experience one fundamental aspect that a modern digital marketplace should address is to safeguard onboarding datasets and services, not only in terms of security but also in terms of legal compliance.

Another key issue that a digital marketplace should provide so as to appeal big organizations such as Banks would be the support of an end2end data/services purchasing process which could be accomplished online. This one-stop-shop approach should include easy to use query mechanisms for seeking data and services that the Bank is interested in, a flexible billing system, which will be able to support different purchasing types (e.g., subscriptions, pay for an individual service/dataset), as well as a legitimate online contract which secures access rights and IPR management.

The Bank is also interested in finding ways to have access to up to date information streams about new trends regarding financial data and fintech services, as well as, the capability to get a preview of this information before purchasing it.

Although the Bank is a systemic bank in Greece, thus it is bound by national regulations and legislation as well as ECB's regulations regarding the financial data, which are processed; it finds interesting the idea of providing part of those data, if a trusted platform existed. Such a platform should be able to guarantee personal data protection, data anonymization, privacy preservation, GDPR and legal protection.

Finally, a major aspect for the Bank is that of security, thus safe user management, strong authentication mechanisms, and unimpeachable audit logging is pre-requisite for the Bank in order to trust a digital marketplace.

## Annex IV: DMA Insights

**Question 1: Description of current process of onboarding applications/services and needs for improvement (e.g. application compliance rules/standards, testing process/period, SLA, contract negotiation between the application/service owner and the data marketplace, smart contracts, life cycle, T&Cs for using the application, data type used, etc.)**

In the DMA prototype, new services were deployed on DMA's OpenShift platform. OpenShift, a platform based on Kubernetes, standardizes interactions between services, provide a scalable infrastructure. The first step of service deployment was the creation of a new OpenShift application. Subsequently, the application was connected to the remote Git repository where the code of the service resided. The establishment of a Git webhook allowed automatic updates of the service whenever new code was pushed to the repository.

Services had their own metadata description, which had the following five categories:

- General service properties (20 items)
- Technical properties (5 items)
- Performance properties (5 items)
- Security properties (2 items)
- Rating properties (1 item)

**Room for improvement:** It has turned out that OpenShift is complex and hard to handle. The choice for OpenShift was partially based on the requirements of a previous project partner. Kubernetes or Docker Swarm, provide similar functionality and should be considered as alternatives.

**Question 2: Description of current process of onboarding data/metadata and needs for improvement (e.g. SLA, contract negotiation between the data/meta data owner and the data marketplace, smart contracts, life cycle, T&Cs for using the data, descriptors, standards, etc.)**

The DMA prototype allowed onboarding of data via a wizard for smaller datasets and via an HTTPS upload for large datasets. The onboarding procedure of the wizard had the following five steps:

- Step 1: Definition of a label serving as a named identifier.
- Step 2: Provision of title and description
- Step 3: Add additional information, such as the e-mail addresses of contact points and publisher, theme, language, and price model
- Step 4: In this step, the dataset can be uploaded, for example as a .csv file. To double-check the uploaded dataset, DMA showed a menu with the provided directory structure.
- Step 5: If the upload was successful, the last step was the publication of the actual dataset

**Room for improvement:** from a usability perspective, the separation of labeling the data and providing title/description is cumbersome. Condensing these two items into one might prove helpful for an improved user experience.

In the prototype, data publishers and customers negotiated contracts with the system "agreed", which is a blockchain-based system in the form of Ethereum smart contracts. The procedure started with the creation of an offer by the data publisher. Interested customers would have been able to provide an offer. In case the offer



**Schema Mapper**

Step 1 **Step 2** Step 3

**Step 2: Match Fields**

Some help text to explain this

Title ↔ Type or select

Description ↔ Name

Publisher ↔ Two

Tags ↔ Type or select

Language ↔ Two

Contact ↔ Type or select

Access Rights ↔ Type or select

< Previous Next >

DMA's search engine let users search for both services and datasets. It was integrated into DMA's overall architecture via an API. Search results could be narrowed down to search facets, which further helped users find their desired services and datasets. The figure below is a screenshot of the search interface.

Portal

Startseite Dashboard Circles (Beta) Suche DMA Labs Hilfe Sprache

**Suche**  
für Datenservices und funktionale Dienste

Sorted By Publisher

Text: Klimareferenz

Typ: Service, Datenkatalog, ☒ Datensatz, Datenverteilung, Datenkreis

Anfangsdatum: yyyy-mm-dd

Enddatum: yyyy-mm-dd

Suchen

1081 results found for the selected filters.

Datensatz | ID: 85653w37-6a23-4d54-845d-07b172ca061  
**Lufttemperatur - Mittel Jahr (Jänner-Dezember), Jänner und Juli - Klimareferenzkarte 1961-1990 Österreich**  
Interpolierte 30 jährige mittlere Klimadaten aus den Jahren 1961 bis 1990. Basierend auf geprüften Messdaten aus dem Stationsmessnetz der ZAMG...

MEHR ZEIGEN

Datensatz | ID: 247a204a-5a78-4390-9999-bec0f0a0c03b  
**Niederschlagssumme - mittlere Summe Sommer (April-September), Winter (Oktober-März) - Klimareferenzkarte 1961-1990 Österreich**  
Interpolierte 30 jährige mittlere Klimadaten aus den Jahren 1961 bis 1990. Basierend auf geprüften Messdaten aus dem Stationsmessnetz der ZAMG...

MEHR ZEIGEN

Datensatz | ID: e46d14a3-5ae1-46e4-b1b8-36f3aa776d1a  
**Lufttemperatur - Mittel Jahr (Jänner-Dezember), Jänner und Juli - Klimareferenzkarte 1961-1990 Österreich**  
Interpolierte 30 jährige mittlere Klimadaten aus den Jahren 1961 bis 1990. Basierend auf geprüften Messdaten aus dem Stationsmessnetz der ZAMG...

MEHR ZEIGEN

Bestitzer: zamg  
Favorit hinzufügen  
Item einblenden

Bestitzer: zamg  
Favorit hinzufügen  
Item einblenden

Bestitzer: zamg  
Favorit hinzufügen  
Item einblenden

In addition to the search interface, where users can explicitly search for desired content, DMA's prototype also had a recommendation engine. This engine suggested relevant services and datasets to users based on their profiles and previous transactions.

#### Question 4: Data marketplace compensation process (e.g. how the fee for using a service/dataset is allocated to the service/dataset owner and the marketplace operator, etc.)

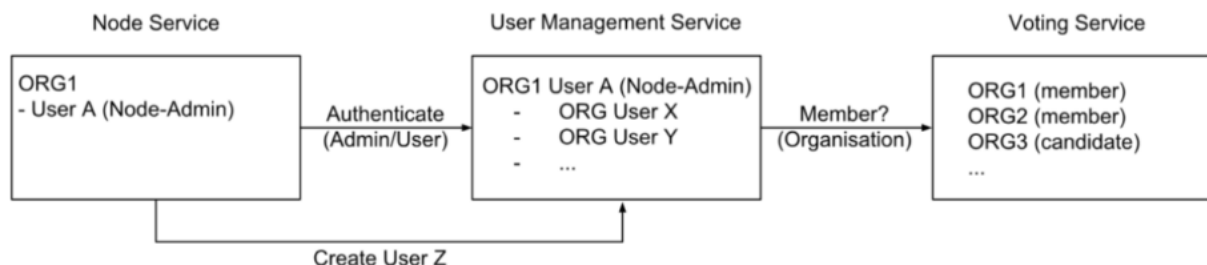
Payment was accomplished via the billing engine, which was supposed to generate invoices based on data, service, and infrastructure consumption. The price of datasets was set during the upload and publication procedure. Users searching for a dataset were able to filter by price model and sort by price.

#### Question 5: User/Subscriber enrolment process (roles, etc.)

The subscription of new members, i.e. data providers, was an election process that needed to be confirmed by existing members. DMA membership was typically targeted at organizations. Members could, but did not have to run their own blockchain nodes.

As mentioned, new members underwent a voting procedure. The procedure started with the registration of a new candidate, which triggered the voting service and started a continuous polling. Once an agreement had been reached, the new member was considered as accepted. The last step was role management, where new members were assigned the status of node admins or node managers.

The figure below gives insights about the voting procedure. Member nodes could only create new members if the voting service indicated them as accepted.



Node admins had the right to create new user accounts, and could add data or services. Node managers were decision makers, for example when voting for new candidates.

#### Question 6: Marketing activities/expected market growth/niche vs mass market

Currently, there are no marketing activities in progress.

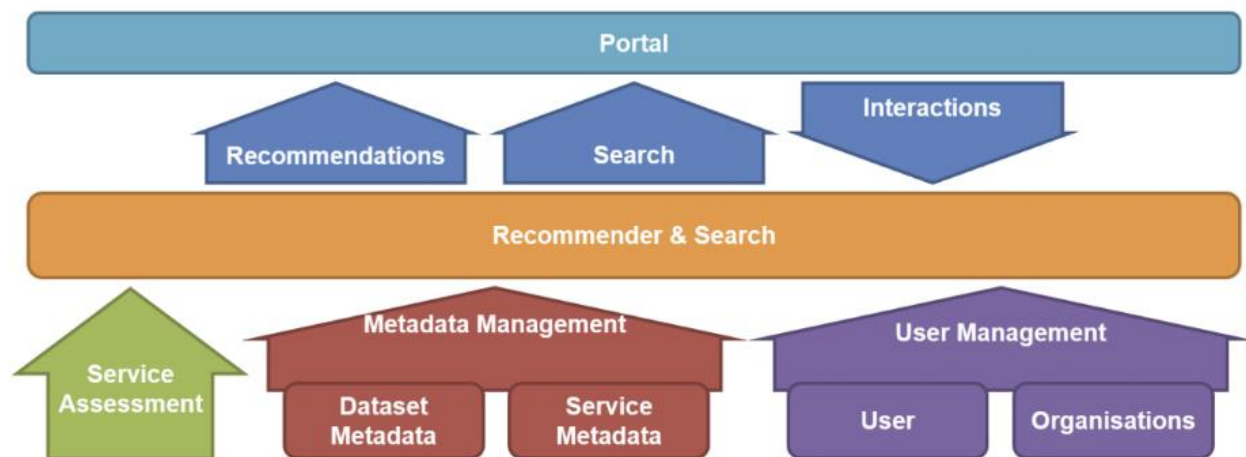
#### Question 7: Roles in operations

Currently, there are no roles defined.

#### Question 8: Infrastructure/architecture/availability/SLA maintenance

The basis of the prototype's architecture were components for metadata management, user management and service assessment. Metadata management covered the metadata for both datasets and services. User management was responsible for users and organizations. Both management systems were connected to the recommender engine and the search engine. The recommender engine suggested datasets and services to users based on their profiles and transaction history. The recommendation engine was not triggered by the

user, instead it was permanently running in the background and delivered suggestions automatically. The search engine helped users to find datasets by using keywords. Faceted search narrowed down search results and gave users fine-grained control over their search. The user-facing component of DMA was the portal. Users interacted with the recommender engine and search engine via the portal. A general overview of DMA's architecture is provided in the image below.



#### Question 9: Other

-