TRUSTS
Trusted Secure Data Sharing Space

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

Authors: **Ioannis Markopoulos (NOVA)**

Additional Information: Report, Public

December 2021

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No 87148

TRUSTS

'Trusted Secure Data Sharing Space'

D2.3 Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II

Document Summary Information

Grant Agreement No	871481	**Acronym**	TRUSTS
Full Title	TRUSTS Trusted Secure Data Sharing Space		
Start Date	01/01/2020	**Duration**	36 months
Project URL	https://trusts-data.eu/		
Deliverable	D2.3 Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II		
Work Package	WP2		
Contractual due date	31-12-2021	**Actual submission date**	20-12-2021
Nature	Report	**Dissemination Level**	Public
Lead Beneficiary	NOVA		
Responsible Author	Ioannis Markopoulos		
Contributions from	Konstantinos Theodoropoulos (NOVA), George Margetis (FORTH), Gianna Avgousti (EBOS), Manos Paschalakis (REL), Evangelos Kotsifakos (LST)		

Revision history (including peer reviewing & quality control)

Version	Issue Date	% Complete	Changes	Contributor(s)
v0.1	02/09/2021	5%	Initial Deliverable Structure and draft ToC	Ioannis Markopoulos (NOVA)
V0.2	17/10/2021	15%	First draft	Ioannis Markopoulos (NOVA)
V0.3	22/10/2021	50%	Include all results from surveys & interviews	Ioannis Markopoulos (NOVA)
V0.4	19/11/2021	90%	Second draft	Ioannis Markopoulos (NOVA)
V0.41	26/11/2021	92%	Review	Stefan Gindl (RSA)
V0.42	26/11/2021	94%	Review	Natalia Simon (IDSA)
V0.43	26/11/2021	96%	Review	Gianna Avgousti (eBOS)
V0.5	20/12/2021	100%	Final	Ioannis Markopoulos (NOVA)

Disclaimer

Copyright message

Table of Contents

List of Figures

List of Tables

Glossary of terms and abbreviations used

Abbreviation / Term	Description
AI	Artificial Intelligence
AML	Anti-Money Laundering
CLI	Command Line Interface
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
GUI	Graphical User Interface
IR	Interview Requirements
KPI	Key Performance Indicator
KYC	Know Your Customer
LR	Legal/Regulatory Requirements
ML	Machine Learning
MPC	Multi Party Computation
NDA	Non-Disclosure Agreement
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 on the work performed in the context of Task 2.2: Industry-specific functional requirements elicitation and analysis (telecom, financial, corporate/personal data).

It is the follow up report of the preceding D2.2 deliverable, which aims at defining the initial TRUSTS functional requirements. In order to define the final set of requirements, the following sources were analysed:

- Updated survey targeting stakeholders of the evolving data marketplace ecosystem.
- Interviews with respective industry representatives.
- Definition and analysis of the EU and worldwide data market trends and industrial needs for growth (D2.1 deliverable)
- Analysis on the state-of-the-art business processes and models (D7.1 deliverable)
- Supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship (D7.4 deliverable)
- Platform architectural requirements (D2.6 deliverable)
- Cycle 1 Trials outcome of the three TRUSTS Use Cases

The analysis of the above sources indicated needed changes to the functional requirements of the deliverable D2.2. In particular, six new functional requirements were added in the categories:

- Purchasing transactions and billing
- Data as a Service and Subscribers management

Moreover, nine out of the 44 functional requirements of the deliverable D2.2 were modified. The current deliverable provides the final set of the TRUSTS platform requirements.

1 Introduction

In today's interconnected world, business silos seem to fade, enabling business expansion to other domains and collaboration of varying industries e.g. by sharing or correlating to other companies' data. Currently however, the creation of a versatile European Data Market and Data Economy is hampered by the lack of trusted and secure platforms, as well as privacy-aware analytics methods for secure sharing of personal data and proprietary/commercial/industrial data. TRUSTS addresses the issue by developing a platform that aims to serve three roles:

a) To create a fully operational and GDPR-compliant European Data Marketplace for personal and non-personal related data.

b) To lay the groundwork for an ecosystem that will enable federation of independent data marketplaces.

c) To become a data ecosystem facilitator, contributing to realization of the expected impact of the HORIZON 2020 ICT-13 Work Programme.

This deliverable constitutes the update of the report D2.2 "Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition I", which contained a detailed analysis of the initial TRUSTS platform requirements and the high level use case scenarios.

Towards advancing T2.2 activities, the implementation of the D2.3 deliverable identifies additional stakeholders and refine questionnaires for the collection of requirements on the TRUSTS platform.

Furthermore, D2.3 incorporates additional information acquired through TRUSTS deliverables and assesses the first cycle of the UC trials lessons learned.

The main goal of the deliverable is to constitute the final set of the TRUSTS platform requirements.

1.1 Mapping Projects' Outputs

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

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,	*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*	*Section 3*	*Section 3 provides an overview of selective current data marketplaces initiatives that are analysed in D2.1 based on respective DMA data collection.*

financial, corporate/personal data)	*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:* *(a) systematic compilation of current data marketplace initiatives, industry related needs, features and capabilities as well as regulatory trends, legislation and standardisation,* *(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. The requirements' capture will involve the documentation of industrial and regulatory needs and opinions about new innovative data 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 NOVA, and through dedicated workshops with the foreseen actors of every case study, animated by NOVA and hosted by every case study owner.*	*Section 4*	*Section 4 provides an overview of Business model requirements that are analyzed in D7.1.*
		Section 5	*Section 5 provides an overview of Intellectual Property Framework requirements that are analyzed in D7.4*
		Section 6	Section 6 *refers to functional requirements analysed in D2.6 referring to the Architecture design of the TRUSTS platform.*
		Section 7	*Section 7 summarizes the functional requirements from the Circle 1 Trials that are reported in D5.10.*
		Section 8 *Section 9*	*Sections 8 and 9 the questionnaires and the interviews with respective stakeholders including telecommunication and banking sectors are referred.*
		Section 10	*In Section 10 the updated list of functional requirements for TRUSTS is provided, juxtaposed to the initial list reported in 2.2, in order to drive platform implementation as well as to be used through the T1.3 methodology for the evaluation of the TRUSTS environment.*
TRUSTS Deliverable			
D2.3 : Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition II [24] *Second 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.*			

1.2 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 compile an updated version of functional requirements, by expanding previous electronic survey and interviews, by juxtaposing previous results to new information from the project WPs and by taking into consideration the initial TRUSTS platform UCs implementation.

In particular, in Section 2 the Methodology is analysed and the requirement sources are defined. Section 3 summarizes industrial needs and requirements. Section 4 refers to business model requirements. Section 5 compiles requirements from analysis of the Intellectual Property framework. Section 6 refers to functional requirements drawn from the TRUSTS platform architectural analysis. Section 7 reports on the initial Trials outcomes. In Section 8 the questionnaires and interview methodology are presented. In Section 9, the survey responses are qualitatively and quantitatively analysed, and the respective requirements are depicted. Moreover, the interviews that are selected from members of the corresponding technology and business communities are analyzed.

Finally, Section 10 lists the resulting updated set of the functional requirements (FRs) and Section 11 depicts concluding remarks on this deliverable.

2 Methodology

The context which governs the TRUSTS requirements elicitation is set by the project objectives. Specifically, TRUSTS Objective 1 (the first of the 7 interdisciplinary objectives as set in the GA is: "To analyse the EU & worldwide challenges and trends for data-sharing and define the requirements for the provision of a multi, concurrent and cross-domain, secure and scalable end-to-end (E2E) data marketplace service".

This objective requires capturing and eliciting end-user requirements, as well as a detailed analysis of the end-user's needs in view of transforming these into specific functional requirements.

To achieve this objective, TRUSTS follows a user centred approach, meaning that the technological design and development obey the stakeholders' expectations. Consequently, the project follows activities outlined by the Ergonomics of Human System Interaction standard (ISO 9241-210[1]), which is part of the multi-part standard ISO 9241 and a revision of the withdrawn ISO 13407:1999. These are:

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 has already produced the first version of the deliverable "D2.2 Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition I", focusing on the two first activities presented above. These deal with the collection, analysis and specification of a first set of requirements for the TRUSTS environment. D2.3 reports on the FRs produced within the T2.2 context. The sources used for the requirements elicitation and analysis are:

1. Updated electronic survey.
2. Domain stakeholders' interviews.
3. Analysis of the EU and worldwide data market trends and industrial needs for growth (D2.1).
4. Analysis on the state-of-the-art business processes and models (D7.1).
5. The analysis of related legal framework, the supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship (D7.4).
6. Architectural considerations for the Use Cases implementation and Business Model realization (D2.6).
7. Cycle 1 trials outcome.

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 above mentioned requirements drive the definition of the updated set of FRs, which will be used for the implementation of the TRUSTS platform, as well as the evolving operational processes design.
- To assist implementation, each FR is mapped to the respective project task.

[1] https://www.iso.org/standard/77520.html

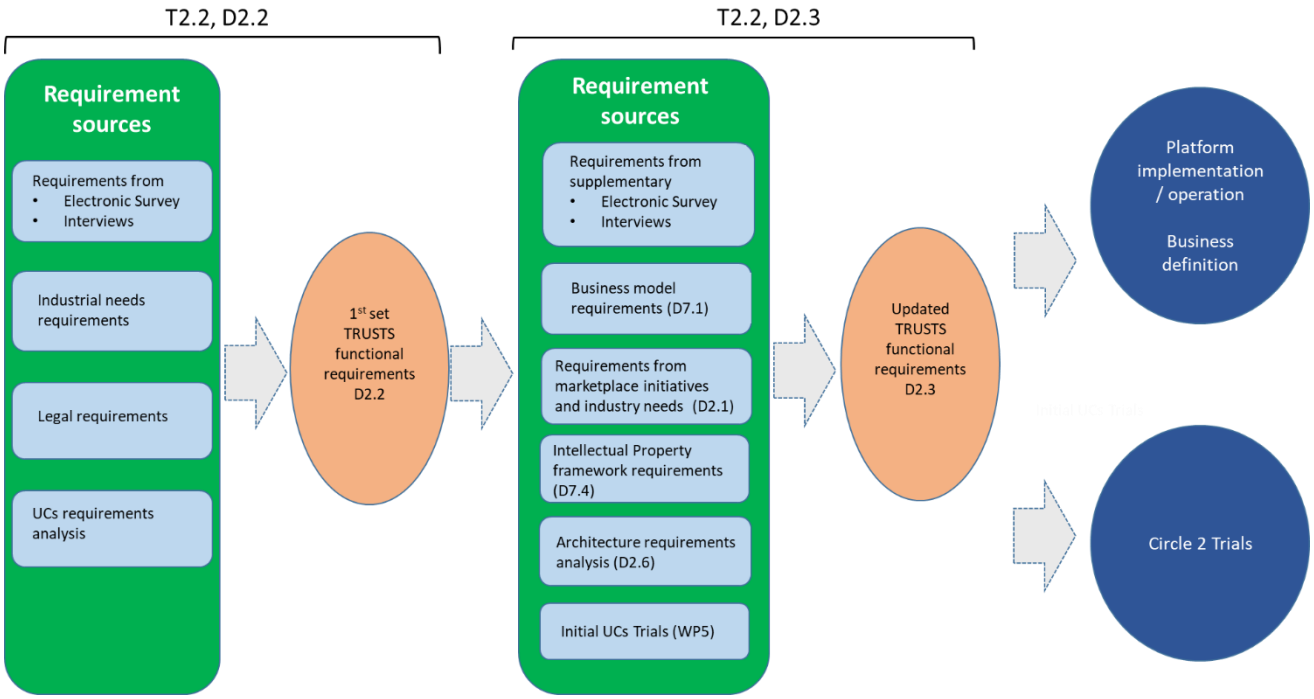

Figure 1: Methodology for the requirements elicitation, analysis and usage.

3 Current data marketplace initiatives and industry needs requirements

This section refers to the state-of-the-art in the data marketplace. The deliverable D2.1 'Definition and analysis of the EU and worldwide data market trends and industrial needs for growth' extracts requirements for the TRUSTS project examining several topics. These are reviewed below.

First, the **analysis of current academic material** mounts into a definition for data marketplaces (relevant even beyond TRUSTS) that point to functionalities a data marketplace should consider having and what challenges data marketplaces are currently facing. A first set of 35 features that are most relevant to data marketplaces are categorised under the market orientation, independent ownership, and many-to-many matching mechanisms are shown in Figure 2.

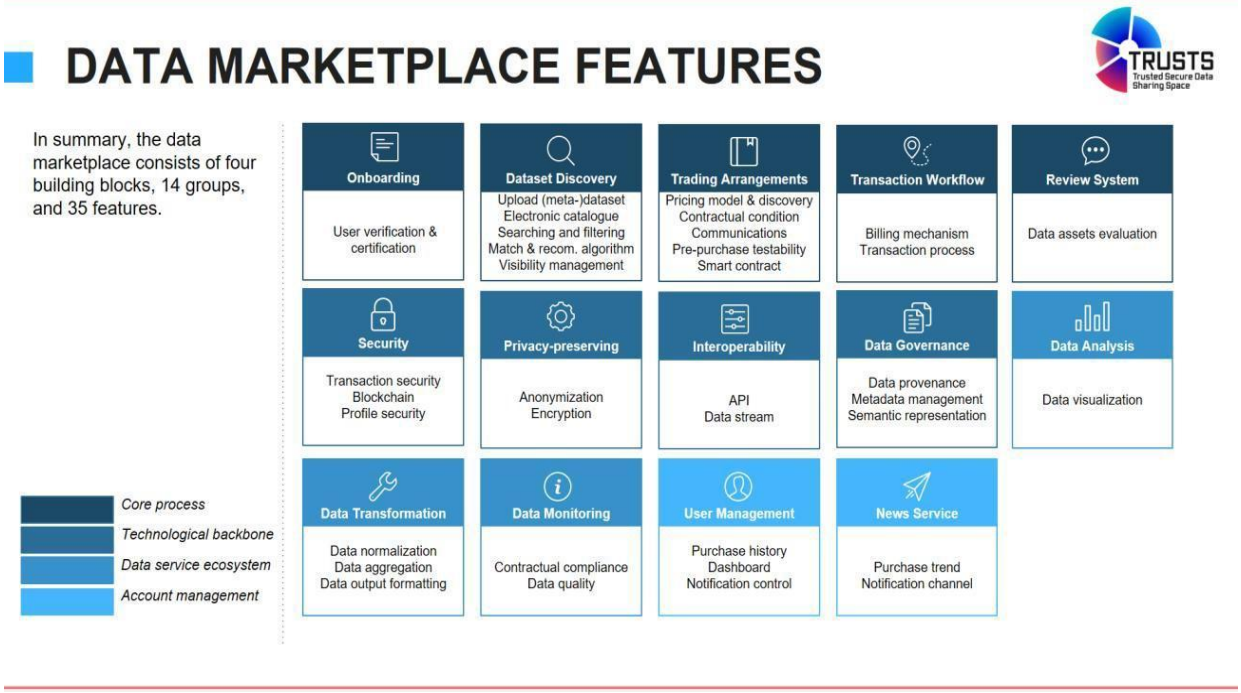

Figure 2: Data marketplace features (from D2.1)

Besides that, the analysis in D2.1 considered the framework of **data marketplaces' macro-environment**, comprising five different areas data marketplaces are influenced by (political, economic, social, technical, legal, and environmental). Several functional and non-functional recommendations for TRUSTS were put forward.

Political considerations suggest that TRUSTS is recommended to:

- be in line with both the EU data strategy and GAIA-X

- connect with the BDVA initiative to establish a vibrant community around TRUSTS

Economic, social, and legal considerations suggest that TRUSTS is recommended to:

- Provide all necessary means to complete the data valuation chain, i.e., either by built-in services or space for third-party apps.
- Enhance trust combating the fears of potential data market customers, i.e., privacy concerns, fear of disclosure of trade secrets, lack of trust in the technical procedures of the platform.
- Implement a persuasive promotion strategy to inform stakeholders about TRUSTS trustworthiness.
- Launch an information and awareness campaign to convince organisations about demand for their data and consequently increase their willingness to share.
- Develop a clear legal framework to guide organisations through the process of developing appropriate business models, as well as the setup of the technology stack required to make their data exchange ready.

Technical developments recommend that:

- TRUSTS should look at the Semantic Data Fabric idea, where the concept of active metadata as well as ML-augmented data ingestion is key and could provide a clear USP and value for the TRUSTS marketplace.
- TRUSTS follow a federated architecture design as well as smart contracting
- TRUSTS should be built on top of existing standards and/or standards under current development like DCAT-AP or IDS/GAIA-X.

Focusing on the **competitive micro-environment** of data marketplaces, it becomes clear that the future TRUSTS operator should

- explore a range of auxiliary services to create a steadfast business moat against the competitive landscape. These include (1) public data harvesting and preparation, (2) support for data provider onboarding through data integration and orchestration, and meta data quality assurance, (3) enablement of co-creation of orchestrated data sets through 3rd party Data Circles as introduced in the Data Market Austria[2] project.
- seamlessly internalise essential services, namely the commissioning/brokerage of the computing infrastructure to facilitate SMEs and data-driven start-ups.

Financial analysis suggests that TRUSTS should

- seek collaboration with players in the financial sector
- implement appropriate APIs towards payment and other financial services, which would facilitate collaboration with financial institutions.

[2] https://datamarket.at/, last accessed Nov 26, 2021.

Considering the **Telecommunications domain**, it is recommended that when commercialisation actions are defined within TRUSTS,

- collaboration with telecom companies should be explored regarding their support for E2E platform operation and additional value creation.

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

MR1	Functionality required by the industry

Remarks

TRUSTS should explore a range of auxiliary services to create a steadfast business moat against the competitive landscape. These include:

- public data harvesting and preparation,
- support for data provider onboarding through data integration and orchestration, and metadata quality assurance,
- enablement of co-creation of orchestrated data sets through 3rd party Data Circles as introduced in the Data Market Austria (DMA) project.

MR2	Data usage processes and services

Remarks

TRUSTS should provide:

- Models for Semantic Representation, where the concept of active metadata as well as ML-augmented data ingestion is key and could provide a clear USP and value for the TRUSTS marketplace.
- Metadata management to inform data origin, content, collection rights.
- Data provenance mechanisms that track data from its origins to destination.

MR3	Services required by the industry

Remarks

- Seamlessly internalise essential services, namely the commissioning/brokerage of the computing infrastructure to facilitate SMEs and data-driven start-ups.
- Create contractual conditions regarding the data (license or contract), for example, data ownership and data usage.
- Implement appropriate APIs towards payment and other financial services, which would facilitate collaboration with financial institutions.
- Data assets should be accessed prior to purchase to evaluate the value of the data (pre-purchase testability).

MR4	Business model requirements

Remarks

- In line with EU data strategy, build TRUSTS on top of existing standards and/or standards under current development like DCAT-AP or IDS/GAIA-X.
- Connect with the BDVA initiative to establish a vibrant community around TRUSTS
- Enhance trust combating the fears of potential data market customers, regarding privacy concerns, fear of disclosure of trade secrets and lack of trust in the technical procedures of the platform.
- Collaborate with telecom companies for support in the end-to-end platform operation and for additional value creation.

4 Business model requirements

This section focuses on the recommendations for the TRUSTS project that are derived from the development of business model taxonomies, which would subsequently inform the design of sustainable business models for TRUSTS, as are laid out in D7.1 'Sustainable business model for TRUSTS data marketplace I'.

A significant highlight of D7.1 is to emphasize TRUSTS' roles in the EU data economy, which go beyond that of a 'basic' data marketplace. TRUSTS will also be a federator and an ecosystem facilitator of data marketplaces. Thus, the business model taxonomies were developed considering these roles, and the recommendations to TRUSTS were organized accordingly. In particular:

- TRUSTS has to fulfill a sustainable role in the data marketplace ecosystem. The aim is to define a multidisciplinary business offering, which is able to address current and emerging needs of commercial enterprises, governmental agencies, academia and individuals while respecting regulations for data privacy, sovereignty and free flow.
- The TRUSTS platform will act independently and as a platform federator, while investigating the legal and ethical aspects that apply on the entire data valorization chain, from data providers to consumers. It will both provide technological means and business incentive to achieve sustainable business and technology federation with third party marketplaces.
- For TRUSTS to become an ecosystem facilitator, it is required to create a business and commercial plan on defining a series of actions that enable data governance models and other framework conditions allowing companies and individuals to avoid the negative externalities of proprietary industrial platforms (supply-driven approach, lower level of control on proprietary data, centralized data governance and technical architecture). Attracting an ever-increasing number of companies and achieving critical mass would be fundamental for TRUSTS to become recognized and successful and a wide range of domain actors.

Analysis of the current status showed market opportunities for TRUSTS, taking advantage of which requires the following:

BR1	TRUSTS business model requirements

Remarks

- It is required that TRUSTS develops a concrete business model aiming at providing targeted solutions for industries, SMEs, professionals and individuals with respect to their data exchange, analysis and trading needs.
- A set of consistent business processes of the platform should be designed aiming at ensuring quality operations, undisputable transactions, IPRs respect and adherence to regulations e.g. GDPR.
- The business model should be supported by a remuneration model aiming at providing fair remuneration of all actors in the value chain. The pricing model should reflect the unique position we envisage for TRUSTS in the data marketplace ecosystem while ensuring clientele loyalty.
- The TRUSTS business model and commercialization roadmap should define the federation model that clearly provides transparency and mutual business benefits with third parties, while ensuring quality of operation.

5 Intellectual Property framework requirements

This section summarizes the results of the deliverable D7.4 "Supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship I". The purpose of this deliverable is to set up the guidelines on how IPR will be managed by the TRUSTS consortium, issuing technical measures to be considered by TRUSTS. These will be further updated and accompanied by contractual measures by M36 in D7.5 "Supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship II".

In particular, the recommended mechanisms for IP protection issued in D7.4 are:

IR1	Data services for IP protection

Remarks

- Securing the IP both physically and digitally by the use of cryptography.
- Data anonymization by removing personal or confidential data fields before publication.
- Federated learning as a technique for decentralized learning where private and sensitive data never have to leave their local storage location.
- Ensemble learning, where aggregated data sets are used.

IR2	Mechanisms to support the IPR protection offered in the International Data Space (IDS)

Remarks

TRUSTS may use mechanisms of the International Data Space, on which TRUSTS is building on, offering the following mechanisms to support the IPR protection:

- The IDS Metadata Broker as an intermediate relevant for TRUSTS IP mapping, as well as providing access and usage control. It functions not only as a search and find function, but also as a gatekeeper that prevents prohibited access to index metadata and prevents the improper use of metadata.
- The IDS Clearing House as a monitoring instance for transactions and indicator for fair use. It has the functionality to track and to monitor that IPR is being protected, as a decentralized and independent service that logs transactions and specific conditions/usage policies, under which data is allowed to be shared.

6 Architecture requirements analysis

In this section we refer to the report D2.6: "Architecture design and technical specifications document I". The TRUSTS platform builds on the experiences and best practices from two previous initiatives for supporting data markets. These are the Data Market Austria (DMA) and the International Data Spaces (IDS). In the Annex I: Architectural Requirements all architecture requirements are shown, labelled ARS 1- ARS 23 and grouped in two categories, depending on whether they relate to functional or technical requirements.

In D2.6 the following D2.2 FR modifications are proposed:

F2R11	The system should ensure the integrity and authenticity of the smart contract transactions signed by its users

Remarks

Previous definition (FR11): "The system should ensure the integrity and authenticity of the smart contracts signed by its users."

Explanation: The reformulation of FR11 is to increase semantic accuracy by encompassing the transaction operation.

F2R12	Smart contracts in the system should be accompanied by a human friendly representation (i.e. natural language)

Remarks

Previous definition (FR12): "The system should provide a human friendly representation of smart

contracts (e.g., natural language)."

Explanation: The current description could imply the automatic derivation of human-readable text from executable smart contract code.

F2R13	Mechanisms to make signed smart contracts be legally valid, enforceable and interpretable will be investigated

Remarks

Previous definition (FR13): "Signed smart contracts should be legally valid, enforceable and interpretable."

Explanation: The current form seems infeasible due to the necessary procurement of legal contracts to underpin the smart contracts.

F2R14	The system should encompass mechanisms for keeping performed transactions from being infringed

Remarks

Previous definition (FR14): "The system should encompass mechanisms for keeping transactions performed ensuring that they cannot be infringed."

Explanation: The new version rearranges the words to clarify the meaning of the sentence.

F2R15	The system should provide the ability to connect to billing mechanisms for enabling consumers to pay providers according to the agreed smart contract

Remarks

Previous definition (FR15): "The system should provide billing mechanisms for enabling consumers to pay providers according to the agreed smart contract."

Explanation: The current wording suggests that TRUSTS will provide a component which implements a payment mechanism. Instead, as reflected by the new wording, TRUSTS will provide a component to connect to existing payment mechanisms.

7 Trials outcome analysis

Cycle 1 Trials provided useful outcome regarding the current TRUSTS platform version (i.e. MVPv.1) operation. Trial outcomes are detailed in the deliverable D5.10 "Performance evaluation and lessons learnt report".

7.1 Use Case Title and objectives

In the current state of development, the TRUSTS marketplace environment is still non-operational as an integrated platform.

The functionalities that were tested are:

- Application/Service onboarding,
- Companies' registration,
- Metadata uploading,
- Asset catalogue search,
- Services execution.

7.2 Use Case Title and objectives

The lessons learned related to usability functional requirements suggest that users demand for:

- fully functional UI (User Interface) to access all TRUSTS functionalities,
- advanced asset catalogue search capabilities, and
- provision of detailed documentation on the TRUSTS usage.

In particular:

Due to the lack of the TRUSTS operational environment, the User Interface tested during the trials was characterized as not adequate in accomplishing basic tasks and in preventing from committing errors. The users opted for a more clean and minimalistic graphical interface with hassle-free sign-up (possibly drag-and-drop) and a business-wise rationale. They also asked for personalized services provided through the UI (e.g. access the personal history of transactions).

The search functionality demonstrated convenient filters during these trials. However, it becomes important to enrich the search process, according to several data attributes like keywords, categories (e.g. field or volume), dataset lifecycle, and targeted recommendations to the customer.

Another issue that was uncovered stems from the trials is related to TRUSTS features providing online help and instructions about marketplace capabilities.

7.3 Use Case Title and objectives

The summarized requirements for the TRUSTS platform indicated by the participating stakeholders in the Trials are listed as TR1-TR9 below.

TR1	**The TRUSTS UI should be easy to use**
Remarks	The use of a Graphical User Interface should be simple. It should be easy to use by both technical and business users.

TR2	The TRUSTS UI should offer services with a business-wise rational
Remarks	The services should include registration, advanced search, buy/sell data, use/provide service, browse data/service catalogue, choose contract, negotiate contract and upload/download datasets.

TR3	The TRUSTS UI should provide personalized information
Remarks	The users demand to be informed about personal trading information, including their own history of the customer's transactions.

TR4	Service/ dataset catalogue
Remarks	TRUSTS should provide selections in the service/dataset catalogue according to several data attributes like keywords, categories (e.g. field or volume), dataset lifecycle, and targeted recommendations to the customer.

TR5	Help function
Remarks	The TRUSTS UI should provide a "Help" function with clear instructions for both technical and business users.

TR6	Presentation of TRUSTS analytics capabilities
Remarks	The users require that the UI provides a more in-depth presentation of the platform's analytics capabilities.

TR7	Adherence to GDPR regulations
Remarks	Focus should be given to the GDPR principles.

TR8	Support Private Set Intersection and Multiparty Computation
Remarks	Private Set Intersection and Multiparty Computation are considered as integral functionalities of the platform and are tested thoroughly.

TR9	TRUSTS platform scalability
Remarks	The platform scalability should be taken into consideration especially with respect to the need to replicate

information in every corporate node.

8 Survey – Questionnaires and Interviews

Following a user-centred approach, TRUSTS analyses the user requirements through a survey and interview of potential end-users of the TRUSTS platform. These include SMEs and service providers who participate in the TRUSTS consortium, as well as extra-project large firms, SMEs and the public sector.

The survey is included in Annex II: Questionnaire and is based on a questionnaire that 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 on selling/buying processes which are followed by the organisation of the survey participant, as well as its intention to use applications and services of a data marketplace in the near future.
- Questions detecting the participant opinion on the standardization gaps and the ways to boost forward the marketplace endeavour.
- Questions regarding services or functions that would make the process of selling and buying data easier.
- Questions about desired characteristics that a data marketplace should be equipped with or avoid.
- Questions with respect to preferred pricing schemes for the use of the TRUSTS platform.

This survey-based data collection has been complemented with qualitative research data obtained through face-to-face interviews, thus allowing to gather more in-depth information. In-depth qualitative interviews with selective relevant interviewees from business sectors related to data markets are undertaken to create a better understanding for the TRUSTS platform requirements, taking into account the various different viewpoints expressed.

The guidelines for the interviews are presented in Annex III: Interviews guidelines

The following section refers to an in-depth analysis of the questionnaire and interview responses.

9 Stakeholders feedback analysis for a commercial use of the TRUSTS platform

In this section, a quantitative and qualitative analysis of the feedback to the questionnaire is reported.

9.1 Electronic Survey

9.1.1 Method and Procedure

The electronic survey was conducted with an electronic questionnaire (see Annex II: Questionnaire), through a GDPR-compliant platform with tools for data privacy disclosures and opt-in statements. The survey link[3] was disseminated to all TRUSTS partners, who were asked to further disseminate it targeting experts in the data marketplace. The survey remained online for more than four months. After a necessary first step for filtering invalid responses, the data was then appropriately grouped and analysed.

9.1.2 Participants

A total of 38 individuals, from 8 countries, responded to the electronic survey (Figure 3). As questions were not mandatory, the total responses per question are sometimes less than the number of respondents. As will be shown in detail below, the majority of participants work in the private sector, which is expected to take advantage as key users of the TRUSTS marketplace. In addition, large organisations that drive the data market as well as private sector SMEs are also well represented in the sample of participants, allowing the survey to record their engagement in data the market and align their expectations with the functional requirements of the TRUSTS platform.

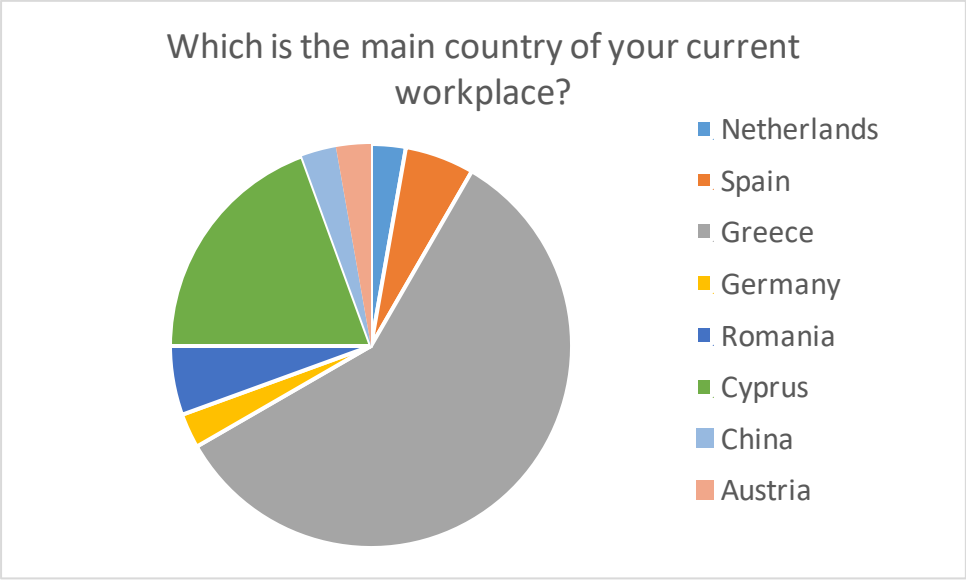

Figure 3: Country of responders' current workplace.

As seen in Figure 4, the majority of participants originate from the private sector, while only 17 % work in academic organisations and another 22 % in the public sector. In addition, most participants work in very large organisations with more than 250 employees (64 %), with 33 % of the responders employed in relatively small firms (1-50 employees).

[3] https://ec.europa.eu/eusurvey/runner/87f209af-2c16-67f0-7794-958e1e8bccf9

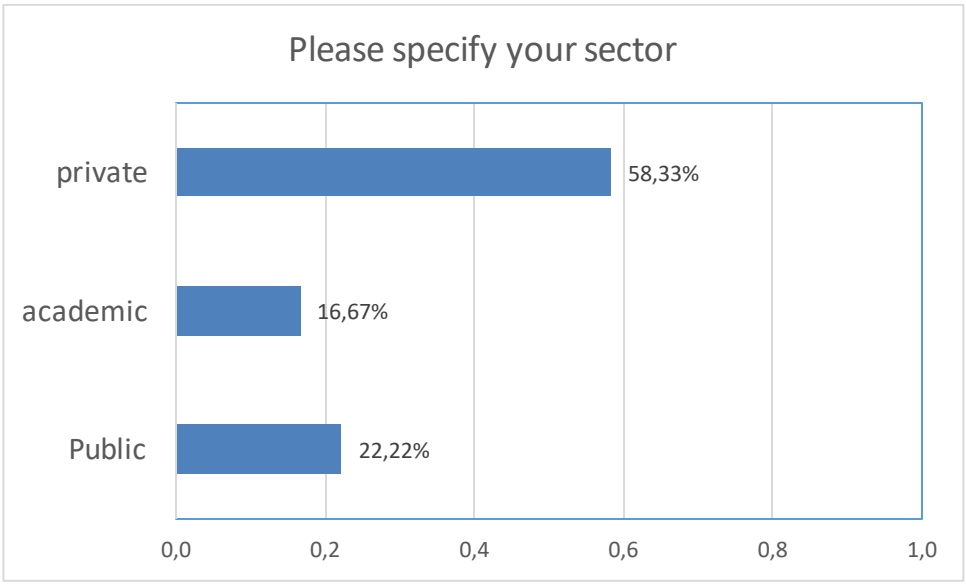

Figure 4: Information on the participants' organization sector.

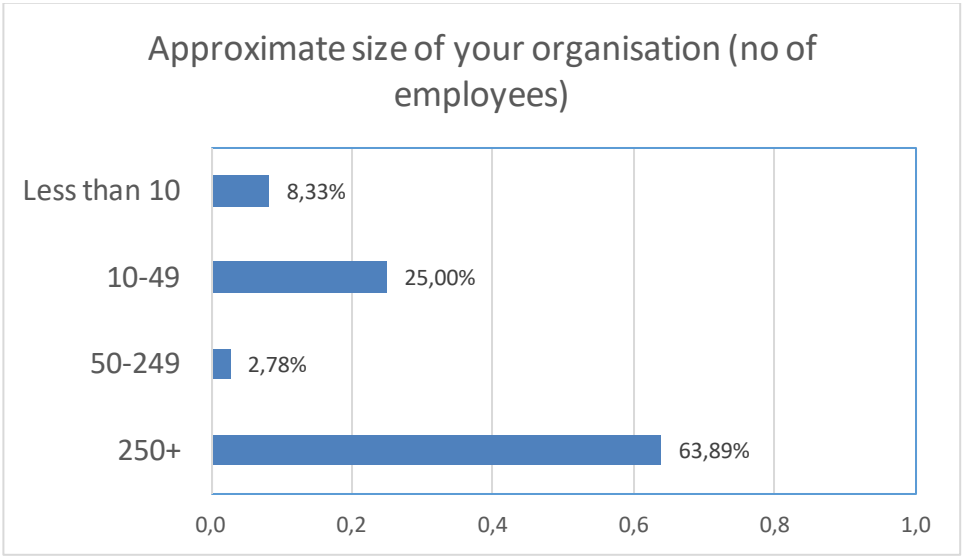

Figure 5: Information on the size of participants' organisations

Another important characteristic of the sample of responders in the survey is that the majority of participants identify themselves as business or technical drivers at all levels of management hierarchy, with many years of experience in the field and a deep understanding of the buying/selling processes in the organisation.

In particular, 44 % characterized themselves generally as business drivers, followed by technical drivers (41 %) and domain experts (33 %). Regarding their level of management, a considerable proportion of participants clarify their role as administrative officers (31 %), closely followed by researchers (28 %). Operating officers and university professors each correspond to 17 %, while executive officers each represent 8 % of the participants (Figure 7).

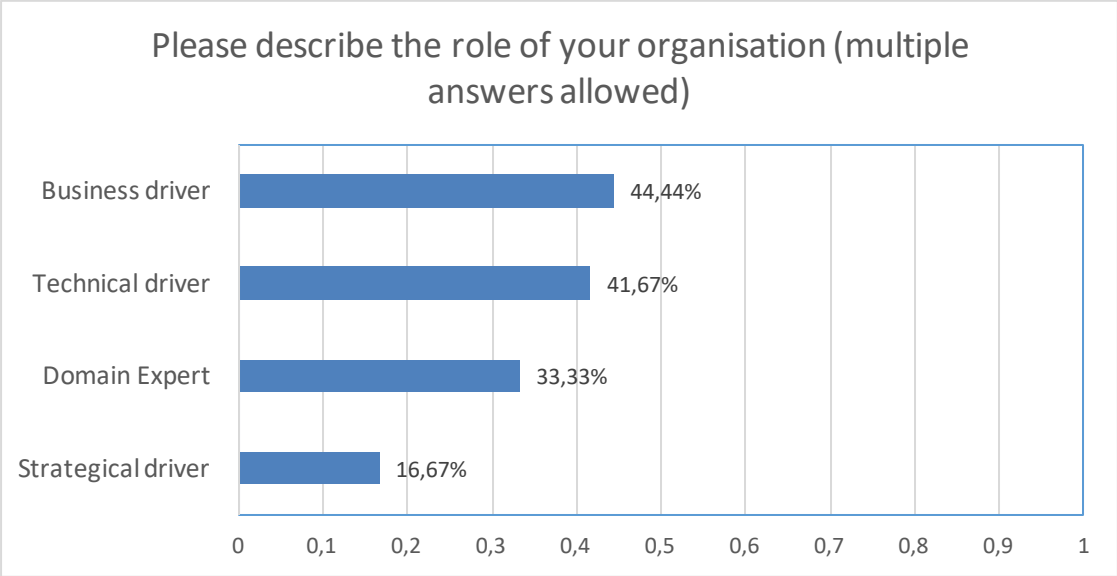

Figure 6: The role of participants in their organisations

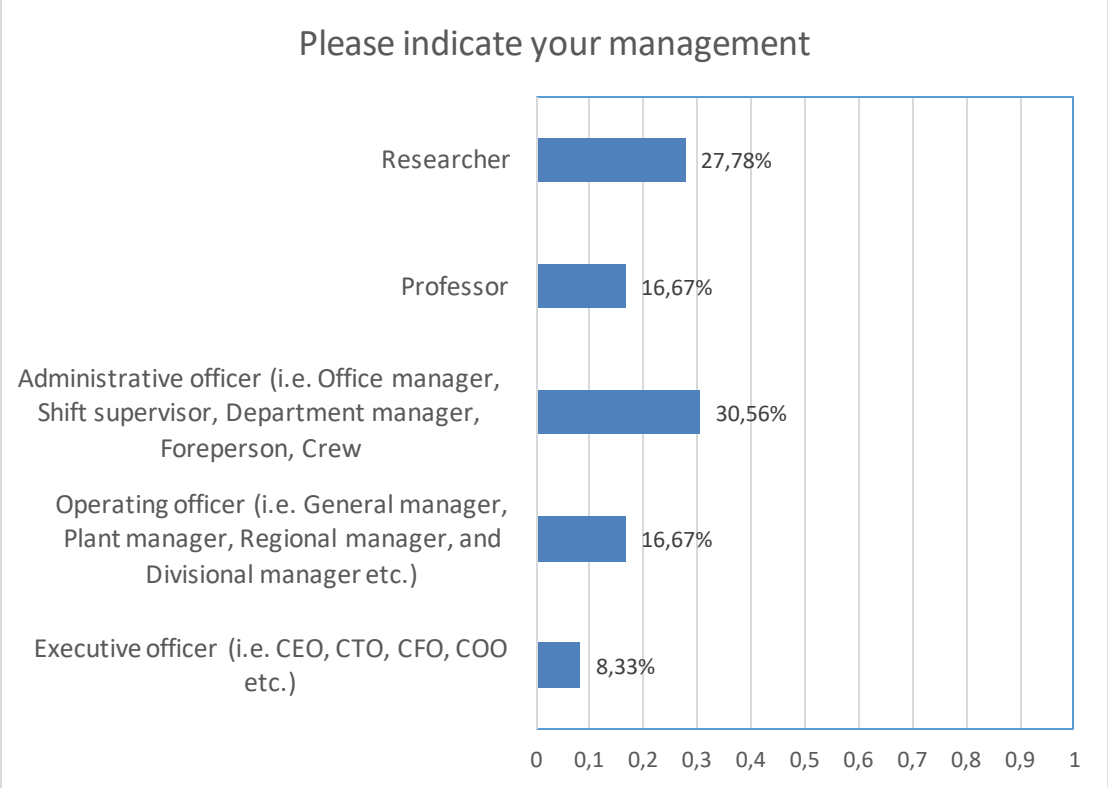

Figure 7: The level of the participants in management hierarchy.

Furthermore, in majority, the participants have many years of business experience, with 58 % in total working in the field for more than 10 years. In detail, 33 % have more than 20 years of experience and 25 % 10 to 20 years. One fourth of the employees has 5-10 years of experience. Three participants reported experience of 2-5 years and another three had less than 2 years in the field (Figure 8). Their experience is also expressed when asked whether they fully understand processes

that involve data training in their organisation. A strong majority 72 % fully or almost fully agreed (4 and 5 on a scale of 5). This is shown in Figure 9.

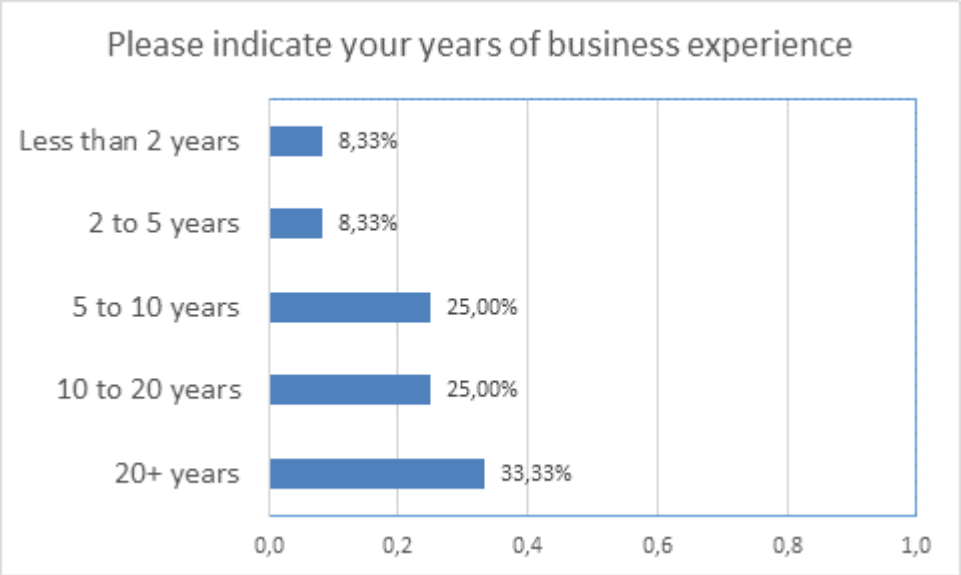

Figure 8*:* Participants' business experience.

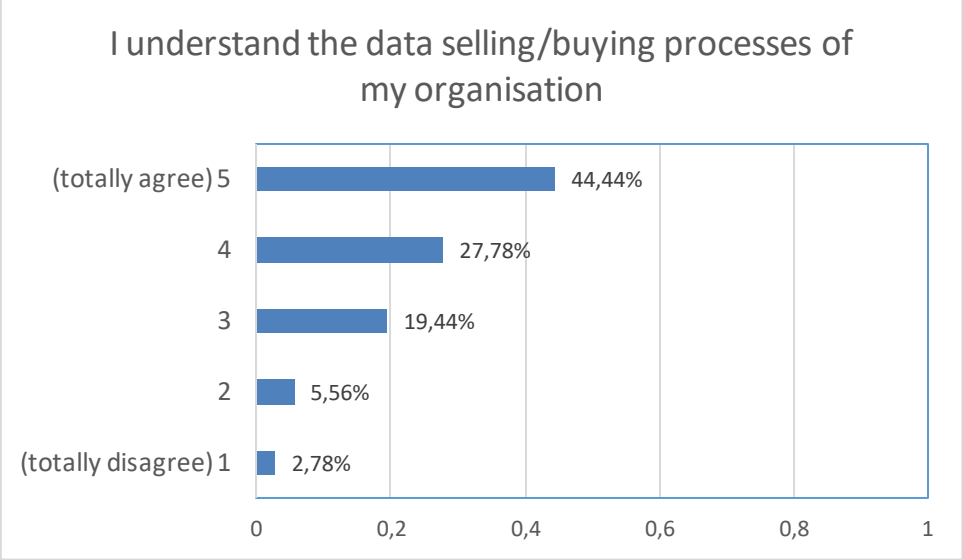

Figure 9*:* Understanding of the processes for data selling/buying.

9.1.3 Users' conception for the data marketplace

The questionnaire explores the conception of the responders for the data marketplace processes they would like to follow. Each participant responded to the question **"Please describe the data marketplace process for providing applications/services you would like to follow"** noting the most important feature of a desired data marketplace. We can group the responses according to this marketplace feature:

- Emphasis on assets: A data space accessed through trading applications. Easy to use for both procuring and selling datasets. Integration to enterprise systems offers reports and data analytics applications. Ensure data quality. Machine generated data maintaining the genuine datasets structure.
- Emphasis to ease of access: A single online store with easily accessed databases filtered by sector, domain, spatial scale. Provides data through flexible APIs or direct processes according to a predefined standard for a certain data domain. Also provides clear-written instructions of the interface to browse, search with a flexible search key, identify sources, compare and purchase data.
- Emphasis on trust: A data marketplace should be a transparent and GDPR compliant ecosystem. Access to be conditional to provided services (requiring analytics to monitor access). Actions should be logged. Data masking service. For research data, compliance is required with the ethical codes of the academic community. Privacy issues (e.g. in results of clinical trials or drug analysis reports).
- Emphasis on seller/client interaction: Trading data and services to liaise with customers. Data ownership may be kept and data usage could be monitored. Data valuation. Announce and trade datasets, analytics application and consulting, accept proposals to create bespoke datasets. Flexible subscription/packages/contracts.
- Emphasis to cloud infrastructure & services: In alignment with the technical standards and policy guidelines of GAIA-X and the International Data Spaces Association (IDSA), allowing its domain-specific specialization in terms of data/trading agreements and flows.

Although we observe a wide variation in the responders' answers considering the main characteristic of a data marketplace, the responses are not mutually divergent. They are rather indicative of the proliferation of the objectives that the TRUSTS platform should eventually meet.

9.1.4 Current activities for buying/selling data applications and services

The questionnaire then investigates the activity of the organisation of the responder in **buying available data**. The frequency of such trading is indicated in Figure 10. A third of the responders (33 %) work in organisations that do not purchase data. Another 50 % of the corresponding organisations make 1-50 transactions per month. More active trading occurs in 17 % of the organisations, with 11% of them purchasing data 51-200 times per month and another 6 % exceeding 200 such transactions per month.

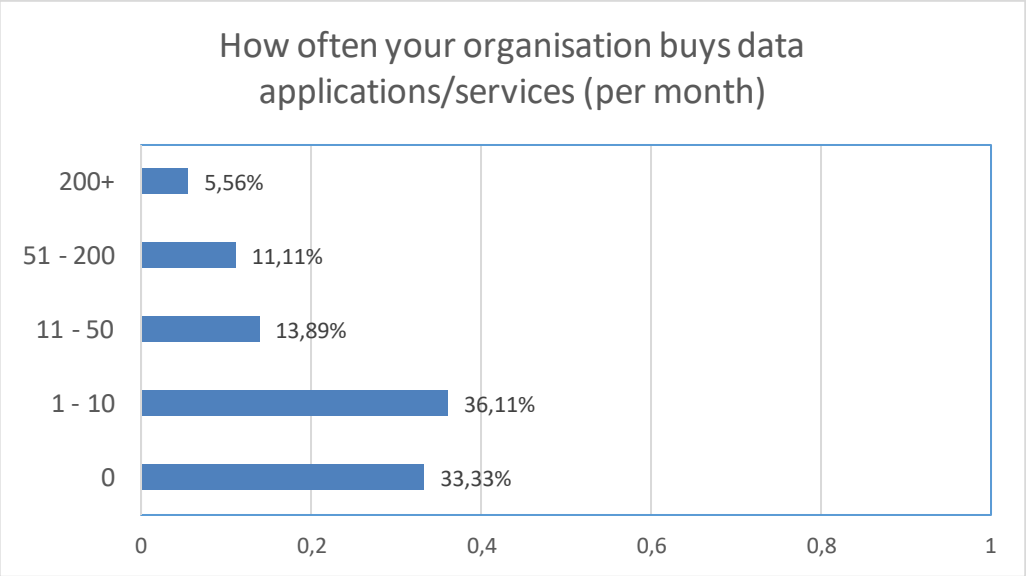

Figure 10: Frequency of purchasing data applications/services

There is noticeable difference in the data applications/services bought, depending on the organisation's size. In particular, the most active organisations (with 50 transactions per month or more) are concerned with:

- Financial and marketing data.
- Public data, governmental data, demographics.
- Research data, open source data.

Less active organisations (less than 50 transactions per month) report buying:

- Reports, surveys, statistical data and projections.
- Market data of candle sticks for financial instruments.
- Load balancing, application performance monitoring, application acceleration.
- Media, lifestyle surveys, CRM.
- Data for research and training AI algorithms.

Regarding the processes to buy applications/services, the responders envision:

- Easily search/discover applications & services, check data quality, preview before buying. Must liaise with consultancy services and applications. Integration with enterprise systems.
- A data space easy to use for both procuring and selling datasets. Negotiation tools and customizable contracts and prices. Traceability of data negotiations and transactions.
- To form data generators with direct long-term contracts
- To have flexible subscription based processes, perhaps yearly or per use subscription if not a constantly used service.
- To facilitate currently followed processes, by procurement of reports based on bilateral agreement with specific providers.

The questionnaire then investigates the activity of the organisation of the responder in **providing data**. Starting from the platforms the organisation currently uses, it asks about the characteristics of a desired future platform providing data applications / services. The frequency of such trading is indicated in Figure 11. Similar to purchasing data, a large percentage of the responders (44 %) works

in organisations that do not provide data. Another 44 % of the corresponding organisations make 1-50 transactions per month. More active data provisioning occurs in 11 % of the organisations, with 5.5 % of them providing data 51-200 times per month and another 5.5 % exceeding 200 such transactions per month.

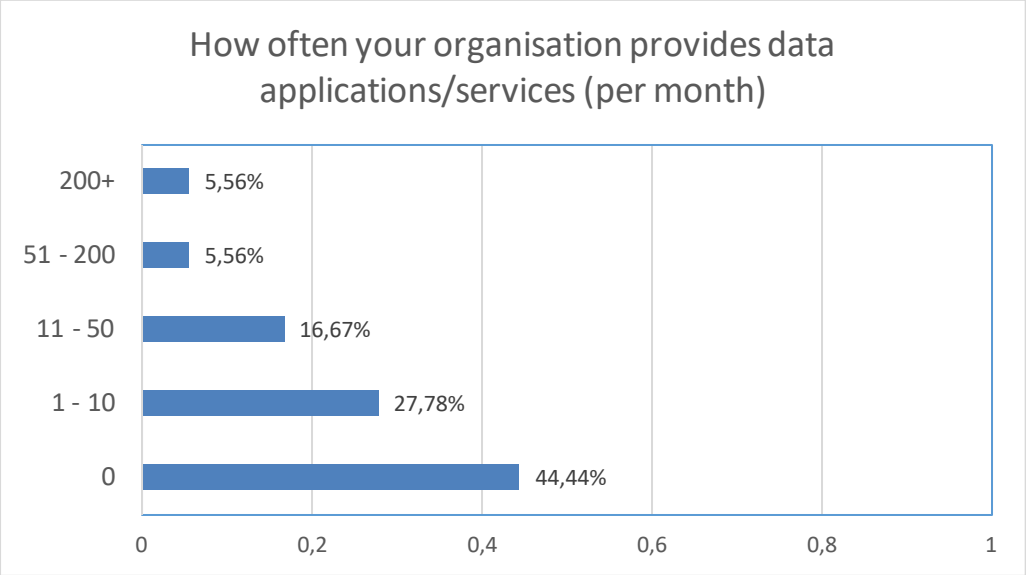

Figure 11: Frequency of providing data applications/services to customers.

In the question "describe the services that the platform you use provides" the responders described their organisation engagement in data applications / service transactions. They referred to:

- Provision of consulting professional services and business analytics
- Financial services for investment management
- Governance, Regulatory, Risk & Compliance Management, Risk intelligence, Scoring & Compliance, Screening, Anti-Money Laundering, etc.
- Data masking services, data sovereignty, identity services
- Collaborative cloud, data exchange and services, repository, git
- Smart city services
- Dissemination of research data that are collected and processed in compliance with the appropriate ethical codes
- Data for research on industry (factories of the future), financial sector, education, etc.

Based on this experience on providing data applications/services, the responders envision a platform that incorporates the following:

- Be able to flexibly integrate the participants' API services. Probably promote them. Include them in search catalogues.
- Be easy to use for both procuring and selling datasets. Create value chains (such applications to match needs are required). Must liaise with multiple dataspaces and data providers. Must liaise with consultancy services and applications. Integration with enterprise systems.
- Offer common spaces and vertical spaces for specific enterprises to avoid having their own trading platform.
- Offer a secure, user friendly interface with precise instructions to integrate services to the platform.

- All transactions should be logged for remuneration and quality reasons.
- Provide a secure channel for easy marketing, reduced communication with the buyer regarding clarifications, simple and prompt announcement of new or updated datasets/services. Data should be updated according to plan.
- Ability to search anonymized datasets of other companies organized by field and volume, preferably functioning as an exchange with other organisations instead of buying and selling.
- Data characteristics and attributes should be exposed to the user prior purchase. Possibly, a test period of sample data could be useful.
- Data valuation would be a significant step to boost data exchange and blockchain could help on this

Provide easy smart contracting facilities and remuneration models to choose. The process should be human friendly for datasets/application owners.

9.1.5 Standardisation

The electronic survey asked respondents to describe the required standardization for federated marketplaces. Besides the general requirements (GDPR, IPR protection, non-repudiation, etc.), and the obvious concern that standards should be followed universally by all participants, the responders pointed to apparent standardization gaps. In particular:

Standardization needs concerning data assets:

- The technologies used in secure data sharing vary. It is not clear how the data should be used, how it should be priced, delivered, etc.
- Data providers should be entitled to securing their data assets and maintain ownership of it.
- Data consumers should know what they are buying. For this purpose, data provenance, collection methods, and data sources are some of the important properties. These properties could be specified in the metadata. Adopt metadata standards.
- Besides the primary users, federated data marketplaces should create the framework for data service providers to make use of the infrastructure for exchanging data. This could be accomplished through decentralization, data segmentation, and providing solutions to the problem with data assets.

Standardization gaps on data exchange and marketplace operation:

- Standardization gaps exist in interoperability, security and privacy, certification mechanisms, as well as in a generic and flexible way of integrating applications and services.
- For a federated data marketplace some minimum industry standards should be in place, facilitating further developments towards federated marketplaces to be easy to communicate and interconnect.
- Data interoperability between IoT and administrative processes and B2B. No common framework for data fusion and anonymization.
- Providers should be able to offer data through standard APIs or known web formats.

There are silos between different technologies and processes in various business sectors. For example, in a bank, a node may be available for Federated learning but not for Multi-party Computing. This makes it difficult to join a marketplace for buying or providing data without extra costs.

9.1.6 Desired services/ functions of the marketplace

Data marketplaces can facilitate the sale- and purchase process of data. Then, the question arises which services or functions are needed from a data marketplace. The results from the current electronic survey are shown in Figure 12.

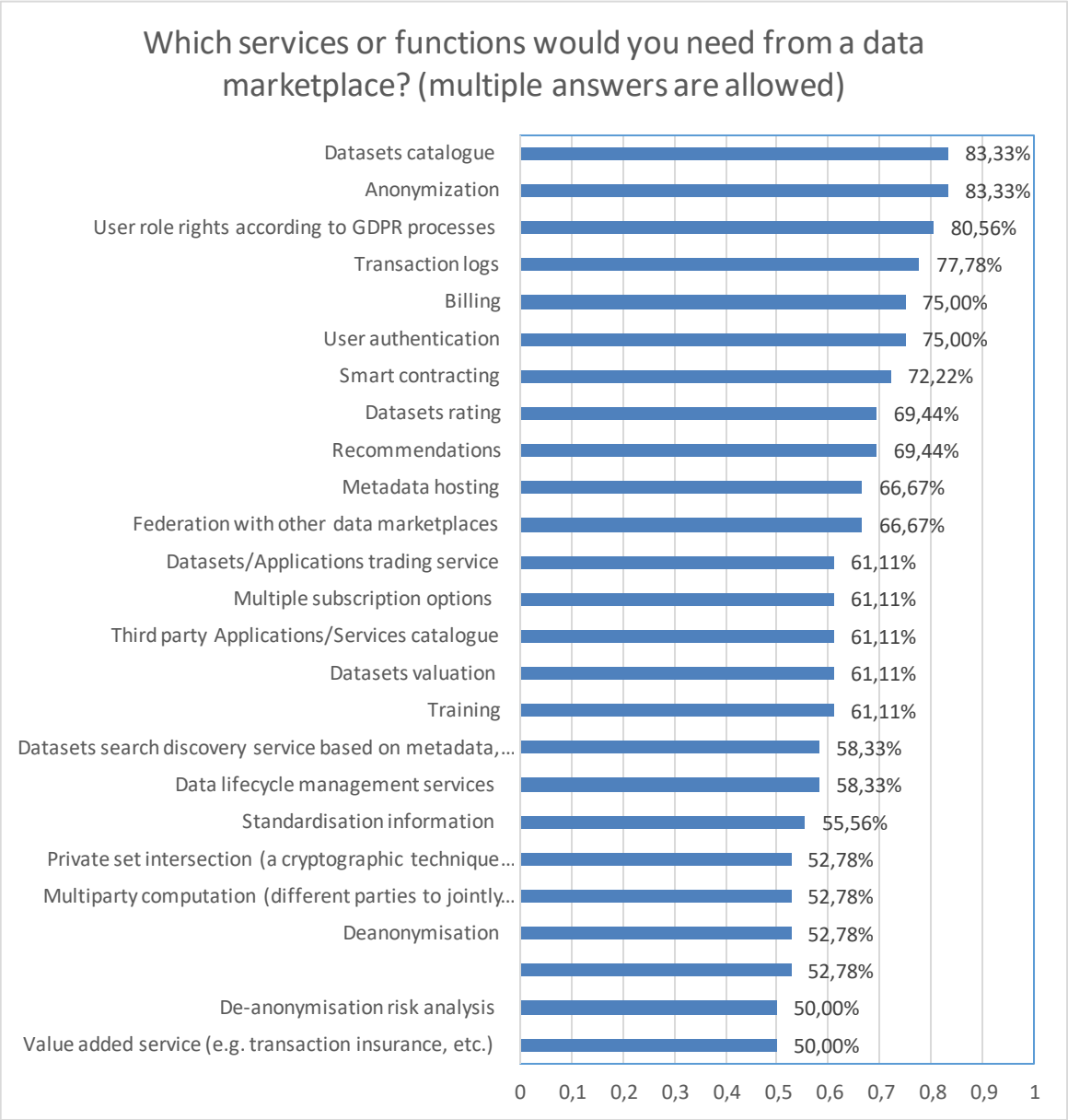

Figure 12: Desired services/functions of a data marketplace.

Below, the results are grouped according to their percentage of positive answers.

Services/functions favored by more than 75% of the responders:

- Datasets catalogue
- Data anonymization
- Transaction logs

- User role rights according to GDPR processes

Services/functions favored by 55 % to 75 % of the responders:

- Billing
- User authentication
- Metadata hosting
- Datasets rating
- Federation with other data marketplaces
- Smart contracting
- Recommendations to the client
- Datasets/Applications towards service
- Standardization information
- Multiple subscription options
- Datasets search discovery service based on metadata, ontologies, etc
- Third party Applications/Services catalogue
- Datasets evaluation
- Data lifecycle management service

Services/functions less favored (though by a strong percentage 50 % to 55 % of the responders):

- 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)
- Multiparty computation (different parties to jointly compute a function over their inputs while keeping those inputs private)
- De-anonymisation
- De-anonymisation risk analysis
- Training
- Transaction assurance and dispute management process
- Value added service (e.g. transaction insurance, etc.)

To the above Services/functions, one should add the following that were proposed by responders:

- Datasets aiming to train AI models.
- Closed vertical or ad hoc groups.
- Service Level Management
- Data visualisation

9.1.7 Innovations and critical points for a data marketplace

The electronic survey responses below focus on the expectations of the participants on advantages / innovations of the platform, as well as on the opposite (i.e., what pitfalls should the platform avoid).

The advantages and possible innovations:

- Unified environment to complement business consultancy services and analytics.
- Single environment for complementing analysis for large organisations with a large number of datasets.
- Unified environment for trading datasets extending internal data analysis processes, with common processes to support interoperability and liability.
- Innovative cross disciplinary environment to create new services and campaigns.
- Faster data trading compared to the traditional agreements between the two parties. Save time for searching and finding required datasets. Companies can spend more time on data processing than data collection.
- Ability to liaise with a great number of dataspaces.
- Reduce the cost of existing data.
- Information security and a robust yet agile data governance with well-defined policies and procedures. Possible desired innovation is the implementation of certification/auditing processes.
- Safer data sharing: Organisations share their data with trust, possibly using blockchain technology. Adoption of a security-by-design approach would increase reliability.
- An innovation could be if the companies can perform calculations and get insights using *encrypted* version of data from all the market participants that are interested
- Create new business opportunities in data sharing. Possibility to commercialize data adding an additional revenue stream in the business models of organisations.
- Facilitate data lineage. Smart data catalogue to define the meaning of critical data elements.
- Provide an AI-driven metadata engine. Data products traded using APIs. Data quality scorecards.
- Data scientists can significantly improve the accuracy and value of predictive models by adding new features to the model.
- Access to new algorithms to be used in analyzing data (expertise in data analytics).

Critical points to be cautious when implementing a data marketplace platform. In particular, the platform:

- Should be able to expand service.
- Should be able to easily correlate or embed external services and applications
- Should not be more complex compared to traditional procuring processes for datasets from individual sources. The whole search, purchase and usage should be archived in a single service.
- Should not initiate operation with a small number of datasets and applications. However, in general, it should emphasize quality over quantity in the content and the service provided.
- Should avoid poor quality, unauthorized, out of date datasets, or data not thoroughly checked for virus. Also it should avoid data overlap or insufficient metadata information.
- Should not overlook GDPR compliance, security (ex, license models) and ethical issues.
- An organisation that builds a data marketplace should consider its role as a data provider with long commitments. Through its business model, it should manage difficulties concerning the ownership and the high cost of third party data in an environment that still lacks a clear legal framework.
- Should caution when using dedicated personnel extensively, as it increases costs and operational complexity.
- Should not focus only on the infrastructure, in expense of establishing the ecosystem of data service providers.

Should not be incompatible or isolated from similar initiatives in the EU.

9.1.8 Pricing models

Regarding the pricing model for the services of the data marketplace, there were no clear favorites (see Figure 13). There were three models, namely the "Fixed price subscription", the "Pay per use" and the "Package" model that were proposed by more than 58% of the participants. The "Free without service level agreement" model and the "Progressive price" model were each voted by approximately 40% of the participants. Moreover, further suggestions were put forward by the responders (labeled "other"). These are:

- Different subscription models for large organisations, SMEs or the Public sector.
- Remuneration models may target SMEs rewarding offering and maintaining datasets.
- All the pricing models should be available. The right pricing model to be chosen depending on the application or after negotiation based on pricing tool proposals.
- Special packages for academic organisations.
- Functioning as a free exchange of datasets between organisations instead of buying and selling.
- Pay after usage according to different contributions for the final result.

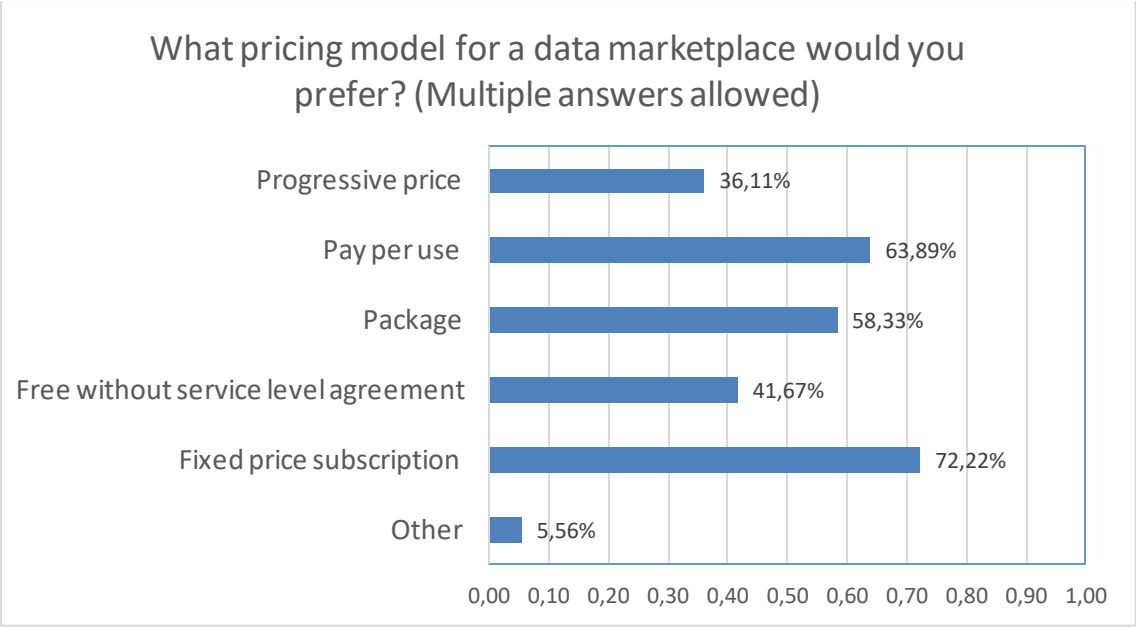

Figure 13: Pricing model options for the data marketplace.

9.1.9 User Interface

An important question of the electronic survey concerned the key characteristics of **the user interface** that a data marketplace should have. The participants in the survey made the following suggestions:

- Drag and drop.
- Clean and minimalistic interface, colorful UI.
- Simple, modern, extremely limited links, no menus.
- A federated data marketplace could be as the google for data (one-page, efficient indexing mechanisms, and output of different providers and their data sources).
- An interface that one would normally expect to see in any e-shop or similar to the product marketplaces.
- Graphical interface with ability to integrate with internal systems.
- Customizability.
- The sign-up process should be quick and hassle-free.
- User profile: wallet icon on the home screen (see the total amount of money that you have collected so far), access your history of transactions (details on what data was sold at what price).
- Search option, data catalogue, categories selection, hot picks, etc.
- Buy/sell data, use/provide service, viewing data/service catalog, choose contract, negotiate contract.
- Register, Upload/download datasets, List of obtained datasets, Help guide.
- Metadata characteristics with data provenance, collection methods, and data sources.
- Marketplaces tools for data retrieval and analytics.
- Community expertise in data management.
- One-click buying.

9.1.10 Barriers for data marketplace usage

The barriers that organisations face using data marketplaces include the following, according to the responders in the survey:

- Lack of a large number of datasets. Multi category datasets are missing.
- Currently a fragmented market. A dominant player is missing.
- There are no current platforms to provide business processes, targeting large enterprises.
- Financial, legal compliances. GDPR compliance, NDAs, and fair IPR treatment is vital.
- Lack of trust and security. Confidence in data sharing is lacking globally.
- Too many hurdles like standardization, certification, costs etc, until you can participate.
- No process for dataset monetization.
- All data sets are visible on a single page (I need to see only the datasets I am requesting – *not all datasets*).
- I don't want to share my data with a third party; I want to be the owner giving access to parties to use my data and decide where to store it.
- Lack of open services and support of open data repositories.
- Data attributes are not accurate, not frequently updated. Samples are not usually available.

- Lack of consistent business models to ensure sustainability. Most data markets close after 2-3 years of operation.
- Lack of federation to access a greater ecosystem.
- Lacking support service, obsolete, outdated, pricy, buggy, incomprehensive.

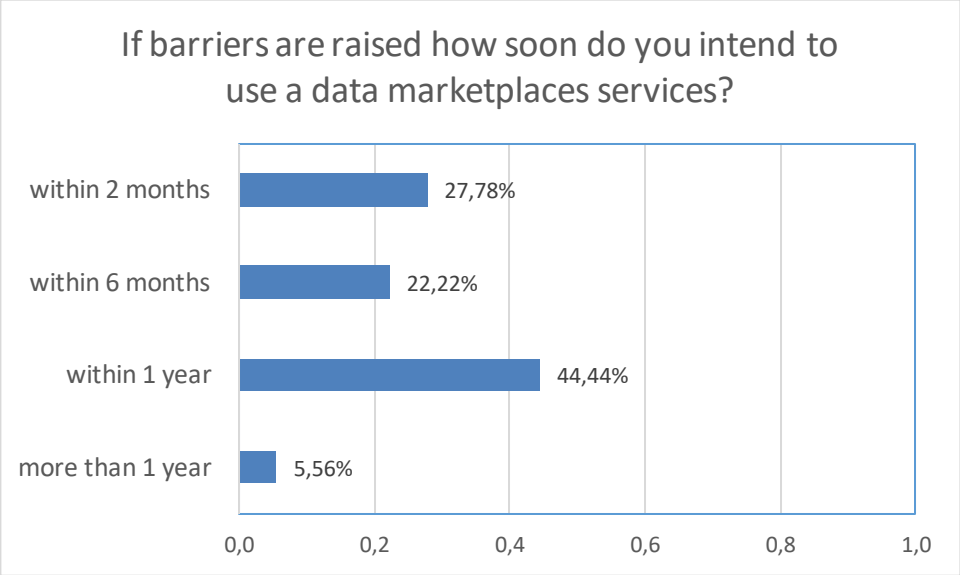

Figure 14: Intention to use data marketplace services.

Once such barriers are raised, the survey participants expressed their intention to use a data marketplace's services as shown in Figure 14. 50 % noted that they would use such services within 6 months after barriers are raised (28 % of which will jump in within only 2 months). On the other hand, 44 % of the responders estimate they can use a data marketplace within a year, while the remaining 6% of the participants think it will take more than a year to do so.

9.1.11 Requirements

This section consolidates the requirements, as they have been identified from the analysis of participants' responses to questionnaires. These requirements typically arise from responses with considerable statistical weight, or stem in similar form from multiple questions. For simplicity, whenever possible, we follow the same method for numbering and defining the requirements that we used in the D2.2.

Same in D2.2	**Q2R1**	**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 throughout the questionnaire responses and, most notably, the specific questions on the conception of the responders for data marketplace processes they would like to follow and their desired user interface. This requirement is the same as QR1 in D2.2.		

Same in D2.2	Q2R2	The process should preserve privacy
Remarks		
This is a main concern of the responders, as a pillar creating trust for using the platform. This requirement is the same as QR2 in D2.2.		

New in D2.3	Q2R3	The platform should support classification of datasets according to their price (paid or free)
Remarks		
This requirement expands Q2R9 to classification of datasets according to price with the capability of on-line search according to this feature. 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.		

Same in D2.2	Q2R4	The E2E platform implementation should provide easy access to the users
Remarks		
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. This requirement is the same as QR4 in D2.2.		

New in D2.3	Q2R5	Data lifecycle management service
Remarks		
This requirement concerns keeping the high quality in the data marketplace, informing buyers about the lifecycle of particular datasets and avoiding out of date ones. It stems from the responses on the desired functions that the marketplace should have.		

Same in D2.2	Q2R6	The process should be accessible upon login on a centralized look up service
Remarks		
This requirement emerged throughout the questionnaire and is directly related to the requirement Q2R4, regarding ease of access to the service and Q2R28 referring to strong authentication mechanisms upon login. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace. It is the same as QR6 in D2.2.		

Modified in D2.3	Q2R7	Searching for datasets should be easy
Remarks		

This requirement was identified mainly when responding to desired user interface characteristics that include search option upon login, data catalogue, categories selection (ex, field or volume), hot pics, etc. Datasets search discovery services should be based on metadata, ontologies, etc. This requirement relates to QR7 in D2.2.

Modified in D2.3	Q2R8 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 Q2R5 and Q2R7. In particular, a federated marketplace could resemble a Google page for data (one page, efficient keyword indexing mechanisms, with output from different providers and data sources). This requirement is related to QR8 in D2.2.

New in D2.3	QR9 Browsing content categories should be supported

Remarks

This requirement was identified in participants' responses to the question regarding the functions that should be provided by a data marketplace, identified as both a datasets catalogue functionality for datasets and a third-party applications / services catalogue. In this case, clear rules for third party application inclusion or federation should be available.

Modified in D2.3	Q2R10 Prior to procuring a dataset, information on the dataset should be provided

Remarks

Data characteristics, attributes, quality scoreboards, ownership and value should be exposed to the user prior to purchase. This requirement also stems from the analysis of participants' responses to the questions regarding the envisioned characteristics of a data marketplace and the functions that should be provided regarding dataset valuation. Related to QR9, QR10, QR21, QR29 of the D2.2. In particular, it was mentioned that value attributes would lead to price comparisons among datasets, with indirect results to facilitate data monetization and eventually bring dataset prices down.

Same in D2.2	Q2R11 The data should have straightforward mapping with the required business entities

Remarks

This requirement (same as QR11 in D2.2) was identified by data buyers, highlighting the need for appropriate data classification and matchmaking with end users' needs.

Modified in D2.3	Q2R12 Processing of personal data should be GDPR compliant

Remarks

Ensure that personal data are processed in compliance to GDPR using, for instance, anonymization protocols. This requirement was specified by data buyers, and is also related to QR18 about GDPR

compliance of the whole buying/ selling process. This requirement also stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace.

Same in D2.2	Q2R13	Efficient processes e.g. one-click payment

Remarks

This requirement for the user interface was identified by data buyers, who asked for a one-click process, highlighting the need for efficient interactions and processes. It is the same as QR13 in D2.2.

Same in D2.2	Q2R14	Provided payment is achieved, direct downloading should be supported

Remarks

This requirement was identified by data buyers, further highlighting the requirement for a one-stop-shop process (Q2R1) and repeating QR14 in D2.2.

New in D2.3	Q2R15	Different subscription schemes should be supported

Remarks

This requirement was identified by data buyers, data sellers and data application/service providers who demand flexibility in subscription methods, including for example fixed price subscriptions, packages, progressive prices, and pay per use. Differentiation between pricing schemes offered to academic institutions, SMEs or big corporations should be also considered.

Modified in D2.3	Q2R16	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, and relates to QR16 in D2.2.

Modified in D2.3	Q2R17	The platform should provide recommendations to the client

Remarks

This requirement was identified by a strong majority of the respondents, who requested recommendations during datasets search.

Same in D2.2	Q2R18	The process should be GDPR compliant and approved

Remarks

This requirement (same as QR18 in D2.2) was identified by data sellers and by data application / service providers, and is also relevant to buyers' concern about data anonymization (Q2R12). 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.

Modified in D2.3	Q2R19	It should be easy to describe the provided datasets through online metadata forms following standards

Remarks

This requirement resurfaced in responders' considerations on the standardization needs for the marketplace and is related to QR19 in D2.2. It 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.

Same in D2.2	Q2R20	The process of data uploading should be easy (one-click)

Remarks

This requirement (same as QR20 in D2.2) was identified by data sellers, as a key characteristic of a convenient user interface for easy online file transfer services.

New in D2.3	Q2R21	Private set intersection

Remarks

This function of the data marketplace is required following the responses in the corresponding question. It refers to a cryptographic technique that allows two parties holding sets to compare encrypted versions of these sets in order to compute the intersection (related to Q2R22).

New in D2.3	Q2R22	Multiparty computation

Remarks

This function of the data marketplace is required following the responses in the corresponding question. It refers to the capacity of two parties holding datasets to compare encrypted versions of these sets, in order to compute the intersection (related to Q2R21).

Same in D2.2	Q2R23	The platform should support direct requests from clients

Remarks

This requirement (same as QR23 in D2.2) was identified by data sellers, as well as by data application / service providers and buyers. It is also related to Q2R24, regarding the direct contact of data sellers with application or service providers.

Same in D2.2	Q2R24	The platform should support targeted selling to specific clients

Remarks

This requirement (same as QR24 in D2.2) was identified by data sellers and is also related to Q2R23. The process as described involved selecting a specific application or service provider, announcing the data to them, and – upon agreements - contracting the provider.

New in D2.3	Q2R25	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.

Same in D2.2	Q2R26	The platform should support open data

Remarks

This requirement (same as QR26 in D2.2) was identified by data sellers and by data application / service providers, but also by data marketplace platform operators.

Same in D2.2	Q2R27	The platform should provide connection to highly visited application marketplaces through federation

Remarks

This requirement (same as QR27 in D2.2) was identified by data application / service providers, who explicitly mentioned as examples of marketplaces Google Play and Apple Store.

New in D2.3	Q2R28	Strong authentication mechanisms should be provided

Remarks

This requirement was identified by participants responding to functions that should be included in the data marketplace. It is related to requirements Q2R4 and Q2R6.

New in D2.3	Q2R29	A dataset rating functionality should be supported

Remarks

This requirement reflects the need for estimating the quality of datasets before purchase. It complements Q2R10.

New in D2.3	Q2R30	Information regarding the anonymization of a dataset should be provided

Remarks

This requirement reflects the need for information provision on the anonymization process while, currently, no common framework for data fusion and anonymization exists. This requirement is related to Q2R12.

New in D2.3	Q2R31	Before purchase, potential clients should be able to have a preview

Remarks

This requirement was identified reflecting their need for estimating the quality of datasets before purchase. It further expands the requirement Q2R10.

New in D2.3	Q2R32	Advanced searching through filters should be supported, including filters related to cost

Remarks

This requirement emphasizes Q2R3 and stems from the analysis of participants' responses to the question regarding the functions that should be provided by a data marketplace, regarding dataset valuation.

New in D2.3	Q2R33	Process for dataset valuation

Remarks

This requirement was identified by sellers who need an estimate for the dataset value. The platform may also have a notification system to tell users when their data can be sold for a high value. The current absence of such a process is characterized in the survey both as a barrier in upscaling the marketplace and as an opportunity for expanding it.

New in D2.3	Q2R34	Transaction assurance and dispute management process

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. Information security with a robust yet agile data governance with well-defined policies and procedures should be put in place.

New in D2.3	Q2R35	The user interface should be simple

Remarks

This requirement stems from the analysis of participants' responses to the question regarding the desired characteristics of the user interface. The responses converge to a clean and minimalistic graphical interface with hassle-free sign-up and drag-and-drop feature.

New in D2.3	Q2R36 The user interface should provide multiple services
Remarks	
This requirement stems from the analysis of participants' responses to the question regarding the desired characteristics of the user interface. The services should include registration, search, buy/sell data, use/provide service, viewing data/service catalog, choose contract, negotiate contract, upload/download datasets. Relates to Q2R1.	

New in D2.3	Q2R37 The user interface should be personalized and informative
Remarks	
This requirement stems from the analysis of participants' responses to the question regarding the desired characteristics of the user interface. Such characteristics include personalization (e.g., wallet icon on the home screen to access the personal history of transactions with details on what data was sold/bought at what price as well as clear help and documentation on the marketplace service capabilities.	

9.2 Interviews

9.2.1 Method and Procedure

In the context of this task, a series of interviews have been conducted targeting experts who use data services. In detail, the TRUSTS partners have performed five one-to-one interviews with employees of a diversity of organisations that had direct or implicit interest in using the TRUSTS platform, following the general guidelines described in Annex III: Interviews guidelines.

9.2.2 Participants

The following Table 2 summarizes the role of each interviewee in the corresponding organisation. Effort was given to interview high-ranking Officers and Managers of the private sector, in order to capitalize on their experience and insight. All interviews are included in Annex IV: Interviews.

Table 2: Interviewees consolidated information

Organisation type	Role in the organisation	Relation with TRUSTS
Large international consulting firm	Executive Officer	Partner

Large international consulting firm	Executive Officer	External
Data creator and service/data selling company in Greece and South Eastern Europe (large enterprise in finance, marketing and business consulting sectors)	Executive Officer	External
Relational Romania	Executive Officer	Partner
Relational Romania	Executive Officer	Partner
IT Development & Business Corporate Services.	Purchase & Sale of data. Data collection from open corporate services providers and analysis.	External
Financial Institution	Banking and investment services. Individual Banking & Business Banking	External
University	Computer Science Professor	External

9.2.3 Requirements

Overall, the interviewees expressed their eagerness for the results of the TRUSTS platform that concern:

- Setting up a fully operational and GDPR-compliant European Data Marketplace, as a one-stop-shop for individual and industrial use.
- Acting independently as a platform federator, allowing the integration and adoption of future platforms.
- Developing consistent business processes designed to ensure quality of operation and undisputable transactions.

The D2.2 deliverable structure has been adopted for the description of the requirements that emerged through the interviews analysis, noting in italics the newly introduced remarks.

I2R1	**Secure and legally compliant exchange of the datasets and services is required.**
Remarks Similar to IR1 in D2.2, 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 *templates with a comprehensive description of the terms of use, facilitating non-disclosure agreements (NDAs).* In addition local laws should apply to each federated node. A suggestion to accelerate business is to provide a set of predefined contracts.

I2R2	**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

(Same as IR2 in D2.2.) 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.

I2R3	**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

Similar to IR3 in D2.2, 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. *The adoption of smart contracts could provide strong business opportunities, given that blockchain-based IoT data marketplaces are still in the conceptual phase.* 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. *Flexibility in contract offerings should be also followed, differentiating between SMEs (which are cost sensitive) to large enterprises (that value certification, stability, help desk, compliance, and clear processes). The latter should also be allowed the possibility to customize the offering.* 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.

I2R4	**Standardised and easy processes should be provided for the onboarding of datasets**

	and services.

Remarks

(Same as IR4 in D2.2.) 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.

I2R5	**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

Similar to IR5 in D2.2, several interviewees stated that 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 metadata should be supported aligned with ontologies and taxonomies of different domains (science, industry, etc). *These should include the case of IoT datasets, which currently lack a common framework for data fusion and anonymization, but are a distinct driver for the proliferation of the data marketplace both in terms of data volume and services.* 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 metadata information, which is kept in the system, needs to be up to date through harvesting external sources on a frequent basis.

I2R6	**Effective and secure user management should be employed.**

Remarks

(Same as IR6 in D2.2.) 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 buyer 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. On the other hand, a seller may profit from a service that provides notifications/recommendations on when his data can be sold for a high

value.

I2R7	Use of intelligent profiling mechanisms for effective matchmaking and recommendation of datasets and services to the users.

Remarks

(Same as IR7 in D2.2.) 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.

I2R8	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

As data volumes continue to explode (e.g. demographics, lifestyle, location, etc.) and machine learning and AI become more important in decision-making, data marketplaces may help organisations reduce the effort and cost involved in locating required data sets and helping data providers extend their market reach. To that end, participants mentioned that AI and ML 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. This requirement relates to IR8 in D2.2.

I2R9	Inherent protection of private datasets should be provided.

Remarks

Similar to IR9 in D2.2, 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, *masking techniques are also extremely useful for securely transfer or store datasets, providing adequate filtering and an extra layer of security.* 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.

I2R10	**Data marketplace should be easy and friendly to use, leveraging productivity and decreasing operational costs through an enriched cost-effective functionality.**

Remarks

(Same as IR10 in D2.2.) 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.

I2R11	Services that allow processing/ customization of datasets.

Remarks

A special need was identified for services over the data, services that will allow the processing of data and their parameterization/customization to fit specific needs, meaning tools that will allow the selection of specific attributes or to perform calculations over the data, e.g. dimensionality reduction, regression etc. A problem would be the interoperability of different data models (coming from different sources), their alignment in order to be able to combine them, maybe through semantic analysis or related services. Another issue would be the ownership of datasets that are produced by applying functions or processing over a dataset. Besides such difficulties, such processing services could provide an important competitive advantage among federated data marketplaces.

9.2.4 Business proposition

During the interviews, participants also referred to strategic and business issues regarding TRUSTS.

In particular, the responders proposed that business may start from vertical markets or big industries as clients and their ecosystem, in order to create dedicated environments for specific industries and wider groups.

They may target a customer base seeking to perform digital transformation (e.g. Fintechs competing traditional banks). Moreover, IoT services have already started to create a great amount of data that may boost the data marketplace businesses as well.

The reputation of the European Union will also serve as an incentive to attract investors/clients/users. Such actors could provide feedback for enriching the platform with datasets and innovative services.

10 Updated functional requirements

10.1 Evolution of functional requirements versus first version

In the preceding sections of this deliverable, various requirements for TRUSTS were noted. Among them, functional requirements are often entangled with business or technical requirements concerning the TRUSTS environment. In this case, the main goal in the Sections 3, 4, 5, 6 and 9 was to extract functional requirements.

10.2 Functional requirements categorization

The updated list of Functional Requirements follows that in D2.2. In particular, FRs are arranged in the following groups:

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

All modified or new FRs are distributed in the above categories.

10.3 Updated functional requirements

Hereinafter the updated functional requirements are presented.

Implementation prioritization has been defined in collaboration with WP3 "TRUSTS Platform implementation":

A. Fully Implementation
B. Partial or Best Effort Implementation

For the detailed analysis on the implementation prioritization please refer to the D3.11 deliverable entitled: "Platform Status Report III".

Table 3: TRUSTS Functional Requirement Specifications

Req. ID	Description	Based on	Task	Priority
Datasets and services onboarding functionality and processes				
FR1	The system should provide standardized API descriptions for enabling providers to onboard their datasets and services	As in D2.2	T3.3, T3.5	A
FR2	The system should provide APIs that enable its interoperability/federation with other industrial marketplaces and external sources	As in D2.2	T3.3	A
FR3	The system should be able to provide datasets and services descriptions	An in D2.2	T3.4	A
FR4	The system should provide reference mechanisms to open data from 3rd sources, so as to make it available as an option through its data exploration, profiling and provision mechanisms.	An in D2.2	T3.3	A
Intelligent data/service exploration and correlation functionality and processes				
FR5	The system should provide rich search mechanisms across all federated nodes for available datasets and services.	An in D2.2	T3.4, T3.5	A
FR6	The system should be able to provide datasets and services recommendations to its' users pertaining to their profile and needs	An in D2.2	T3.6	A
FR7	The system should employ matchmaking mechanisms through which hosted datasets are matched with hosted services (e.g., suitable for their analysis or processing / customization) and vice versa.	An in D2.2	T3.6	A
Modified FR8	The system should identify and match related datasets so as to provide combined and enriched data, with services that allow processing / customization of datasets.	I2R8	T3.6	A
FR9	The system should be able to improve datasets and services profiles based on extracted information originating from the available data	An in D2.2	T3.6	A
Purchasing transactions and billing				
FR10	The system should provide contract mechanisms as a validation means of sellers/buyers agreements	An in D2.2	T3.2	B
Modified FR11	The system should ensure the integrity and authenticity of the smart contracts transactions signed by its users. A dispute management process may be designed.	F2R11, Q2R34	T3.2	B
Modified FR12	Smart contracts in the system should be accompanied by a human friendly representation (i.e natural language)	F2R12	T3.2	B
Modified FR13	Mechanisms to make signed smart contracts be legally valid, enforceable and interpretable will be investigated	F2R13	T3.2	B

Modified FR14	The system should encompass mechanisms for keeping transactions from being infringed	F2R14	T3.2	A
Modified FR15	The system should provide the ability to connect to billing mechanisms for enabling consumers to pay providers according to the agreed smart contract	F2R15	T3.2	B
FR16	The system must provide alternative and flexible pricing models taking into consideration the diversity of the available datasets and services	An in D2.2		B
FR17	The system should provide brokerage mechanisms for addressing the offerings and demands of the hosted datasets and services	An in D2.2	T3.6	A
NFR 1	The system must provide alternative subscription contracts to the TRUSTS platform subscribers.	I2R3		B
NFR 2	The system should provide a set of templates to the asset owners to describe the T&C for using the respective asset through TRUSTS.	I2R1		B
NFR 3	Prior to procuring a dataset, information on the dataset should be provided (including data characteristics, attributes, ownership, etc.) including a sample of the data set for pre-purchased testability and preview	Q2R9, MR3		B
(Meta-)Data Governance				
FR18	The system should provide explicit metadata information for each dataset or service that is accommodated in the platform	An in D2.2	T3.4	A
FR19	The system should incorporate models, ontologies and taxonomies for the classification and semantic representation of the accommodated datasets and services in the platform	An in D2.2	T3.4	A
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	An in D2.2	T3.4	A
FR21	The system should provide mechanisms capable of identifying the provenance of the hosted datasets.	An in D2.2	T3.4	A
FR22	The system should provide mechanisms capable to identify the lifecycle of the hosted datasets	An in D2.2		B
FR23	The system should harvest metadata from external sources.	An in D2.2	T3.4	A
FR24	The system should be able to provide semantic information even for unstructured datasets	An in D2.2	T3.4	A
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	An in D2.2	T3.6	A
FR26	Dataset discovery should be based on the FAIR principle	An in D2.2	T3.4	A
Data as a Service and Subscribers management				

FR27	TRUSTS datasets and services should be provided to the users on demand, regardless of geographic or organisational separation between provider and consumer taking into account all potential territorial legislation/regulatory restrictions.	An in D2.2	T3.5	B
FR28	TRUSTS should be able to be deployed as a federation of distributed, interconnected and interoperable nodes.	An in D2.2	T3.1, T3.5	B
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 seller and the buyer.	An in D2.2	T3.5	A
FR30	The system should support mechanisms for users' (producers/consumers) subscription opting different schemes (e.g., annual, monthly, etc.) and authentication	An in D2.2	T3.5	A
FR31	The system should support corporate accounts that fall under one subscription/enrolment per organisation	An in D2.2	T3.5	B
FR32	The system should enable authorized users to create, read, update, and delete (withdraw or make unavailable) datasets, services and user profile records	An in D2.2	T3.5	A
FR33A	The system should provide validation criteria for the new enrolled users	An in D2.2		B
FR33B	The system should provide reputation/rating schemes with regard to available datasets and services.	An in D2.2		B
FR34	The system should allow consumers to announce their need for specific datasets / services if there are not any available, already.	An in D2.2	T3.5	B
FR35	The system should provide notifications regarding datasets / services updates to users that have granted access to them	An in D2.2	T3.5	A
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.	An in D2.2	T3.5, T5.2, T5.3	A
Modified FR37	TRUSTS UIs and workflows must follow a business-wise rational (e.g., one stop shop), for coherently mapping the market's needs. Indicative services to be included: registration, advanced search, buy/sell data, use/provide service, browse data/service catalogue, choose contract, and upload/download datasets.	Q2R41, TR3	T3.5	A
NFR 4	TRUSTS UIs should be personalized	Q2R42, TR4		B
NFR 5	TRUSTS should provide clear help function and documentation	Q2R37, I2R3, I2R4, TR5, TR6, TR7		B
NFR 6	The trusts platform and service should be scalable.	TR10, AR1		A
Data protection				
Modifie	The system should support collaboration between parties	Q2R21,	T4.1	B

d FR38	while preserving the privacy of the data. Methods that enables data privacy preserving on parties' collaboration will be provided by the system	Q2R22		
FR39	The system should provide de-anonymization attack assessment and risk analysis for the private / sensitive datasets to be onboard	An in D2.2	T4.3	B
Modifie d FR40	The system should employ anonymization tools and guidelines for data anonymization. Information about anonymization of a dataset should be provided upon client request.	Q2R30	T4.3	B
FR41	The system should provide means for converting algorithms that might compromise the data privacy into safe privacy preserving ones without harming their functionality	An in D2.2	T4.5	B
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.	An in D2.2	T4.2	B
FR43	The system must incorporate a secure infrastructure for the distributed analysis of data based on ML approaches	An in D2.2	T4.4	B
Trusted and legitimate data flows				
FR44	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): ● General Data Protection Regulation ● e-Privacy regulation, for electronic communications ● Free Flow of Non-Personal Data Regulation, for data exchange between the TRUSTS platform and subscribers ● Platform-to-Business Regulation, for safeguarding TRUSTS' operational transparency and fairness. Mechanisms provided ensuring that local laws apply to each federated node. Predefined contracts should exist.	An in D2.2	T6.1, T6.2, T6.3, T6.4	A

11 Conclusions

In this deliverable multidisciplinary relevant sources were analysed to derive requirements with respect to the development and operation of a data marketplace. A comprehensive list of Functional Requirements was produced.

This deliverable constitutes the second version of the two reports containing the detailed analysis of the requirements for a commercial financial and operators' industry vertical data marketplace platform. It presents an updated set of functional requirements that was defined in D2.2 entitled "Industry specific requirements analysis, definition of the vertical E2E data marketplace functionality and use cases definition I". This updated list incorporates more recent work in the project WPs, including:

1. The use of an updated electronic survey.
2. Domain stakeholders' interviews.
3. Analysis of the EU and worldwide data market trends and industrial needs for growth (D2.1).
4. Analysis on the state-of-the-art business processes and models (D7.1).
5. Analysis of related supporting mechanisms for Intellectual Property Rights Protection and Data Stewardship (D7.4).
6. Architectural considerations for the Use Cases implementation and Business Model realization (D2.6).

The updated list of functional requirements will be used in the definition of the final TRUSTS platform architecture.

Annex I: Architectural Requirements

The following Table 4 is taken from deliverable D2.6: "Architecture design and technical specifications document I".

Table 4: Summary of Architecture Requirements

ARS ID	Short description	
ARS 1	Smart contracts, which are based on data sharing policies, are supported.	Functional
ARS 2	All transactions on any node, which is part of a TRUSTS platform instance, are logged.	Functional
ARS 3	All data assets / data products in a TRUSTS platform instance, use metadata schema or vocabularies.	Functional
ARS 4	External data marketplaces can be integrated into a TRUSTS platform instance.	Functional
ARS 5	The operator of a TRUSTS platform instance can use an administrative interface. Users of the platform also have an interface available.	Functional
ARS 6	Every TRUSTS platform instance is a distributed set of nodes.	Functional
ARS 7	Each node runs software components, which are distributed as docker containers, and can be configured via configuration files.	Technical
ARS 8	Each node needs to run an instance of the trusted connector. It is required for communication between nodes and within a node	Technical
ARS 9	Each node has an internal directory of running software components	Technical
ARS 10	A node can internally be a distributed node itself.	Technical
ARS 11	Every TRUSTS platform instance has at least one node which runs an instance of the Broker component.	Technical

ARS 12	A node must be able to run apps on premise.	Technical
ARS 13	Assets (such as data assets) in a node can be managed locally.	Technical
ARS 14	A node is able to harvest metadata from its local assets.	Technical
ARS 15	Recommendations in a node are integrated into the TRUSTS platform, which provides data, delivers recommendations and returns feedback.	Technical
ARS 16	Instances of the trusted connector support existing security infrastructure inside of a node, and connect it to the usage control of the TRUSTS platform.	Technical
ARS 17	Computation using distributed and controlled execution environments is possible in every instance of the TRUSTS platform.	Technical
ARS 18	Machine Learning using distributed and privacy preserving technologies is possible in every instance of the TRUSTS platform.	Technical
ARS 19	Every instance of the TRUSTS platform provides the option of execution of distributed and privacy preserving technologies on nodes provided by the operator of the instance.	Technical
ARS 20	Every instance of the TRUSTS platform provides the option of execution of distributed and privacy preserving technologies on servers provided by a TRUSTS participant themselves.	Technical
ARS 21	Development of distributed services using privacy preserving technologies is possible in every instance of the TRUSTS platform. (This refers to Devops.)	Technical
ARS 22	Every instance of the TRUSTS platform provides scalability for services and apps running on the platform instance.	Technical
ARS 23	The user interfaces of all TRUSTS components are integrated into the CKAN extension mechanism.	Technical

Annex II: Questionnaire

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.

In this questionnaire you will be asked about the data sharing processes in your organisation, therefore it aims at people that are having the need to exchange or trade data in your organization.

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 31st 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 is 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**

- I agree
- I do not agree

Requirements Elicitation Questionnaire

Demographic information

*** Please provide the name of your organisation**

*** Please specify your sector**

- Public
- Private
- Academic
- Regulatory body

*** Please indicate the approx. size of your organisation (no of employees)**

- Less than 10 persons employed
- Less than 50 persons employed (10 to 49)
- Less the 250 persons employed (50-249)
- 250 or more persons employed

*** What is the main country of your current workplace?**

- AF - Afghanistan

- AL- Albania

DZ - Algeria

AD - Andorra AO - Angola

AG - Antigua and Barbuda

AR - Argentina

AM - Armenia

AU - Australia

AT - Austria

AZ - Azerbaijan

BS – Bahamas

BH - Bahrain

BD - Bangladesh

BB – Barbados

BY - Belarus

BE - Belgium

BZ – Belize

BJ - Benin

BT - Bhutan BO - Bolivia

BA - Bosnia and Herzegovina

BW - Botswana

BR - Brazil

BN - Brunei Darussalam

BG - Bulgaria

BF - Burkina Faso

BI - Burundi

- CV - Cabo Verde
- KH - Cambodia
- CM - Cameroon
- CA - Canada
- CF - Central African Republic
- TD - Chad
- CL - Chile
- CN - China
- CO - Colombia
- KM - Comoros
- CG - Congo
- CR - Costa Rica
- CI - Côte D'Ivoire
- HR – Croatia
- CU - Cuba
- CY - Cyprus
- CZ - Czechia
- CD - Democratic Republic of the Congo
- DK – Denmark
- DJ - Djibouti
- DM - Dominica
- DO - Dominican Republic
- EC - Ecuador
- EG - Egypt
- SV - El Salvador
- GQ - Equatorial Guinea
- ER - Eritrea

- EE - Estonia
- SZ - Eswatini
- ET - Ethiopia
- FJ - Fiji
- FI - Finland
- FR - France
- GA - Gabon
- GM –Gambia
- GE - Georgia
- DE - Germany
- GH - Ghana
- GR - Greece
- GD - Grenada
- GT - Guatemala
- GN - Guinea
- GW - Guinea Bissau
- GY - Guyana
- HT - Haiti
- HN - Honduras
- HU - Hungary
- IS - Iceland
- IN - India
- ID - Indonesia
- IR - Iran
- IQ - Iraq
- IE - Ireland
- IL - Israel
- IT - Italy
- JM - Jamaica

- JP - Japan
- JO - Jordan
- KZ - Kazakhstan
- KE - Kenya
- KI - Kiribati
- KW - Kuwait
- KG – Kyrgyzstan
- LA - Laos
- LV - Latvia
- LB – Lebanon
- LS - Lesotho
- LR - Liberia
- LY - Libya
- LI - Liechtenstein
- LT - Lithuania
- LU - Luxembourg
- MG - Madagascar
- MW - Malawi
- MY - Malaysia
- MV - Maldives
- ML - Mali
- MT - Malta
- MH - Marshall Islands
- MR - Mauritania
- MU - Mauritius
- MX - Mexico

FM - Micronesia MC - Monaco

MN - Mongolia

ME - Montenegro MA - Morocco

MZ - Mozambique

MM - Myanmar

NA - Namibia

NR - Nauru

NP - Nepal

NL - Netherlands

NZ - New Zealand

NI - Nicaragua

NE - Niger

NG - Nigeria

KP - North Korea

MK - North Macedonia

NO - Norway

OM - Oman

PK - Pakistan

PW - Palau

PA - Panama

PG - Papua New Guinea PY - Paraguay

PE - Peru

PH - Philippines

- PL – Poland
- PT - Portugal
- QA - Qatar
- MD - Republic of Moldova
- RO - Romania
- RU - Russian Federation
- RW - Rwanda
- KN - Saint Kitts and Nevis
- LC - Saint Lucia
- VC - Saint Vincent and the Grenadines
- WS - Samoa
- SM - San Marino
- ST - Sao Tome and Principe
- SA - Saudi Arabia
- SN - Senegal
- RS - Serbia
- SC - Seychelles
- SL - Sierra Leone
- SG - Singapore
- SK - Slovakia
- SI - Slovenia
- SB - Solomon Islands
- SO - Somalia
- ZA - South Africa
- KR - South Korea
- SS - South Sudan
- ES - Spain
- LK - Sri Lanka
- SD - Sudan

- SR - Suriname
- SE - Sweden
- CH - Switzerland
- SY - Syrian Arab Republic
- TJ - Tajikistan
- TZ - Tanzania
- TH - Thailand
- TL - Timor-Leste
- TG - Togo
- TO - Tonga
- TT - Trinidad and Tobago
- TN - Tunisia
- TR - Turkey
- TM - Turkmenistan
- TV - Tuvalu
- UG - Uganda
- UA - Ukraine
- AE - United Arab Emirates
- GB - United Kingdom
- US - United States of America
- UY - Uruguay
- UZ - Uzbekistan
- VU – Vanuatu
- VE - Venezuela
- VN - Viet Nam
- YE - Yemen
- ZM - Zambia

- ZW - Zimbabwe

*** Which of the following roles fit you best?**

- Business driver
- Strategical driver
- Technical driver
- Domain Expert

*** 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

*** Please indicate your years of business experience**

- Less than 2 years
- 2 to 5 years
- 5 to 10 years
- 10 to 20 years
- 20+ years

Data marketplaces are digital environments that enable datasets trade and data applications services usage. Data marketplaces operate under various business models e.g. subscription based, etc. In the forthcoming questions we would be grateful if you could provide your insights for current and envisaged datasets/applications exchange processes.

* **I understand the data selling/buying processes of my organisation (1: totally disagree, 5: totally agree)**

  - **Only values between 1 and 5 are allowed**

* **Please describe the data marketplace process for providing applications/services you would like to follow**

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

  - **Only values between 1 and 5 are allowed**

* **Please describe the data applications/services you buy**

*

* **Please describe the data marketplace process for buying applications/services you would like to follow**

* **Please tell us how often your organisation provides data services through the platform you use (per month)**

* **Please describe the services that the platform you use provides**

* **Please describe the data marketplace process for providing services through the platform you use, that you would like to follow**

* **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**

* **Data marketplaces can help make the process of selling and buying data**

easier. Which services or functions would you need from a data marketplace?

- ☐ Anonymization
- ☐ Deanonymisation
- ☐ De-anonymisation risk analysis
- ☐ Metadata hosting
- ☐ Datasets catalogue
- ☐ Datasets valuation
- ☐ Datasets rating
- ☐ Third party Applications/Services catalogue
- ☐ Datasets search discovery service based on metadata, ontologies, etc.
- ☐ Recommendations
- ☐ Data lifecycle management services
- ☐ 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
  - ☐ Smart contracting
  - ☐ Multiple subscription options
  - ☐ Standardisation information
  - ☐ Training
  - ☐ Transaction assurance and dispute management process
  - ☐ Federation with other data marketplaces
  - ☐ Value added service (e.g. transaction insurance, etc.)
  - ☐ Other

* **Please describe the advantages/innovations of a data marketplace**

* **Please describe what a data marketplace should avoid**

* **What pricing model for a data marketplace would you prefer?**

- Free without service level agreement Fixed price subscription
- Package
- Pay per use
- Progressive price
- Other

* **Please tell us the key characteristics of the user interface that a data marketplace should have**

Please inform us on the barriers you face on using current data marketplaces

*** If barriers are raised how soon do you intend to use a data marketplaces services**

within 2 months

within 6 months

within 1 year

more than 1 year

*** Would you like us to contact you to further discuss your view on the data marketplaces?**

- Yes
- No

Annex III: 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's answers by requesting clarifications or examples when he/she responses for a company process, data exchange, applications needed, etc.

We do not focus only on 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 organisations 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 organisations and employees within these organisations, 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 IV: Interviews

Interview NOVA 1

Means: Face to face

Date: 17/9/2021

Interviewee affiliation:

Large international consulting firm

Key outcome:

With more devices deployed, large organisations are the ones with the greatest incentive to find a way to squeeze more return out of their investment by taking advantage of the IoT marketplaces. IoT services will create a great amount of data that will boost other businesses as well.

Currently, blockchain-based IoT data marketplaces are still in the conceptual phase. However, momentum is building as a growing number of entities make progress in moving their applications forward.

The emerging initiatives are building an ecosystem of hundreds of organisations that are helping to create proof-of-concept marketplaces that can subsequently be scaled. These organisations are creating the conditions within their marketplaces that allow "data traders" to maximize their data asset value.

Three initiatives are especially important:

1. Custom platform technologies: Blockchain technology is evolving.
2. Dedicated services: Marketplace developers are also creating ancillary services to promote data trade.
3. Target customer base: Digital transformation and competition (e.g. Fintechs competing traditional banks) providers to create dedicated environments for specific industries and wider groups.

There is a lot of potential but in order to be successful one has to access sources of really big data respecting security and IPRs.

Business may start from vertical markets or big industries as clients and their ecosystem.

A business alignment with such industrial partners could be beneficial.

With respect to the competition, one may take into account that AMAZON, EBAY and other existing marketplaces may turn into data marketplaces as well. One to succeed has to be quicker, secure, robust and having created a significant ecosystem.

Interview NOVA 2

Means: Online telco

Date: 24/9/2021

Interviewee affiliation:

Large international consulting firm

Key outcome:

In order for a data marketplace to be successful the operator has to convince its clients to run part of their business through the data marketplace.

Fintech and IoT are two distinct drivers for the proliferation of the data marketplace market both in terms of data volume and services.

The payment process if significant to sustain business and increase trust.

A data marketplace product manager had to be flexible.

Need to differentiate offerings addressing large enterprises from SMEs.

Large enterprises are not really interested on small prices but the value certification, stability, help desk, compliance, and clear processes. They also would like to have the possibility to customize the offering. Marketplaces offer opportunities for large enterprises to increase efficiency. To succeed, they must create exceptional experiences that build trust with both sellers and buyers. From seller and buyer onboarding, to payments, to data delivery logistics, personalised online experience.

SMEs are cost sensitive. So respective packages should exist targeting SMEs.

Interview NOVA 3

Means: Online telco

Date: 27/9/2021

Interviewee affiliation:

Data creator and service/data selling company in Greece and South Eastern Europe (large enterprise in finance, marketing and business consulting sectors).

Key outcome:

The company assists, among others, large companies towards their digital transformation offering applications (e.g. marketing tools) and data.

Data are both generated locally (e.g. through surveys, etc.) and procured through international affiliates.

Typical data types for sale in a data marketplace can range from business intelligence and research, demographic, and market data to business intelligence and public data.

A data marketplace is public or commercial form of data sharing.

Data sharing has a long history in academic, research, and public policy circles but has in recent years made enormous progress into private enterprises, from big business to analyst, consulting, and market intelligence firms.

Data consumers include government, analysts, big business, and market intelligence firms.

As data volumes continue to explode (e.g. demographics, lifestyle, location, etc.) and machine learning and AI become more important in decision-making, data marketplaces may help organisations reduce the effort and cost involved in locating required data sets and helping data providers extend their market reach.

This could be a unique selling point for those who operate data marketplace and may attract customers vs the competition. The data marketplace should be a hub to other related services and a unique point of sale and data provision for the clients.

Interview EBOS 1

Location: GoToMeetings

Date: 28/09/2021

Interview affiliation: IT Development & Business Corporate Services.

More than 15 years of experience. Purchase & Sale of data daily. Daily collection of data from open corporate services providers and analysation by company's means.

A data marketplace is seen as an online transactional location or store that facilitates the buying and selling of data.

They want to sell data and to buy data.

The data ecosystem is a complex, fragile network of relationships and stakeholders, and like any strong relationship, these connections require **trust**.

Confidence in data sharing is lacking globally.

With thoughtful applications of emerging technologies like blockchain and privacy enhancing techniques, public and private sector operators of data marketplaces can empower people to grant permission willingly and knowingly for their data to be used.

One must design a robust governance structure that builds trust.

One to succeed must be quicker, secure robust and having created a significant ecosystem.

Some interesting features that would like to have:

- a clean and minimalistic interface and a colourful UI
- the sign-up process should be quick and hassle-free.
- should allow users to have some form of control over which type of data they wish to sell.
- should allow users to see how their data is being used
- should have a notification system to tell users when their data can be sold for a high value.
- should allow users to withdraw their money into their credit cards
- should show the user a rundown of all data that is available of him for each platform.

Interview EBOS 2

Location: Skype

Date: 04/10/2021

Interview affiliation: Financial Institution.

Providing banking and investment services. Individual Banking (checking accounts, savings accounts, debit/credit cards, etc.) Business Banking (business services, checking accounts and savings accounts for businesses, treasury services, etc.)

_____

A data marketplace is a platform on which anybody (or a great number of potentially registered clients or suppliers) can upload and maintain data sets.

A data marketplace is a more public (and sometimes commercial) form of data sharing. A one-stop-shop for buying and selling external data. Access to and use of the data is regulated through varying licensing models.

Data marketplaces are two-sided markets. There's the data provider, who is looking to commercialize their data assets, and there's the data buyer, who wants to find a data source which meets their requirements.

For data buyers, they should make data sourcing effortless.

For data providers, they should maximize data monetisation. Data monetization is a way of companies increasing their revenue by selling their intangible data assets.

A data marketplace is an amazing solution to the problems often associated with accessing, managing, and monetizing external data.

It must comply with relevant legal and ethical regulations when it comes to data collection.

Use of Blockchain is essential to ensure that all data exchanged over the platform are encrypted.

Setting up platforms for selling, buying, needs to clarify:

- What kind of data will be available?
- Who would buy the data?
- How much would they pay for it?

On the inside:

- Identify the data value
- Identify two main users

- Create appropriate control and data governance.
- Set the appropriate licensing models and contracts in place before exposing the data to consumers
- Build functions to monitor, measure and continuously improve data products and features.

Interview Relational Romania 1

Date: 17/09

Means: Skype

Interviewee affiliation:

Relational Romania/*Department of R&D EU Projects*

Please describe the data marketplace process for providing applications/services you would like to follow

We are interested in the benefits of having our Data masked through the Data masking process. We realize how important it is for data that we need to securely transfer or store to have an extra layer of security that the data masking process can well provide.

Please describe the data applications/services you buy

For this specific moment in time we do not buy any data applications or services as we have not found any according to our standards and budget. However, we provide an adequate filtering to our data that we would like to improve in the future.

Please describe the data marketplace process for providing services through the platform you use, that you would like to follow

The data marketplaces process that we envision is one that provides services through the platform. It consists of a secure channel between all the parties that are involved. It has less communication with the buyer regarding any clarifications. Also, it should have a simple and prompt announcement of new datasets/services that are registered in the marketplace. Finally, the marketing should be one of an "easy-to-go" style.

Please describe in your opinion the standardization gaps and the way forward to boost the data marketplace endeavor/Please describe the required standardization for federated data marketplaces

Nothing to mention.

Please describe the advantages/innovations of a data marketplace

A colossal advantage for companies that are participating in a data marketplace of a kind like TRUSTS will be that through the reputation of the European union and the innovation and security of the platform, many investors will be interested in having their part on their platform and as a result many companies will see a huge impact in their economic growth. Additionally, we should not forget about the customer as it will reduce the cost of existing data.

Please describe what a data marketplace should avoid

A big mistake that data marketplaces do is they have a dedicated staff to run the ins and outs of the marketplace. This has a result of increasing heavily the costs of having the marketplace running and it further complicates the smooth operation of the marketplace when and if there is a minor hiccup.

Please tell us the key characteristics of the user interface that a data marketplace should have

In order for a marketplace to strive, its interface needs to be attractive and modern. There should be no complicated menus whatsoever.

Please inform us on the barriers you face on using current data marketplaces

The current barriers that we face are those of legal matter such as GDPR and NDAs

Interview Relational Romania 2

Date: 17/09

Means: Skype

Interviewee affiliation:

Relational Romania/General Manager (GM)

Please describe the data marketplace process for providing applications/services you would like to follow

We're curious in the advantages of having our data filtered and processed using the Data masking procedure. We recognize the importance of having an extra layer of protection for our data, and in the case that we need to securely transfer or retain them, the data masking procedure is essential.

Please describe the data applications/services you buy

We do not buy any data apps or services at this time because we have not identified any that meet our expectations and budget. However, we provide adequate data filtering, which we hope to improve in the future.

Please describe the data marketplace process for providing services through the platform you use, that you would like to follow

The data marketplace approach that we envision is one in which trustworthy services are provided via the platform. It is made up of a secure channel that connects all of the parties involved. It has a lower level of communication with the buyer when it comes to any clarifications. It should also feature a simple and prompt notification of newly registered datasets/services in the marketplace. Finally, the marketing technique should not have any gimmicks and it should be straightforward.

Please describe in your opinion the standardization gaps and the way forward to boost the data marketplace endeavor/Please describe the required standardization for federated data marketplaces

Nothing to mention.

Please describe the advantages/innovations of a data marketplace

Many investors will be interested in participating in a data marketplace like TRUSTS because of the European Union's reputation, as well as the platform's innovation and security. As a result, many enterprises will enjoy a significant increase in their economic growth. Furthermore, we should not overlook the customer, with that being done it will lower the cost of existing data.

Please describe what a data marketplace should avoid

The fact that data marketplaces many times have a specialized team that administer the ins and outs of the marketplace is a significant mistake. As a result, the costs of running the marketplace will increase significantly, and the seamless operation of the marketplace will further be complicated when and if there is a tiny setback.

Please tell us the key characteristics of the user interface that a data marketplace should have

A marketplace's interface must be appealing and current in order for it to succeed. There should be no menus that are overly complicated.

Please inform us on the barriers you face on using current data marketplaces

The present roadblocks are legal in nature, such as GDPR and non-disclosure agreements (NDAs).

Interview LST 1

Means: online meeting

Date: 6/10/2021

Interviewee affiliation:

Computer Science Professor, GR University

Key outcome:

Data marketplaces would be very important for researchers so they can find a variety of data to perform experiments. Currently they are not able to find large datasets and most of the available ones are of bad quality. Data not only for the IT or commercial/industrial community but also for other sciences should be available to help research in other areas as well (for example linguistics), heterogeneous data.

There is the need for well-organized, structured data in order to avoid extra manual processing after acquiring them. Annotated data would be also very useful – or services that will allow an automatic annotation (given specific parameters) of a dataset.

There is also a special need for services over the data, services that will allow the processing of data and their parameterization/customization to fit specific needs, meaning tools that will allow the selection of specific attributes or to perform calculations over the data, e.g. dimensionality reduction, regression etc.

A problem would be the interoperability of different data models (coming from different sources), their alignment in order to be able to combine them, maybe through semantic analysis or related services.

Good search capabilities are required – find the most appropriate datasets to acquire.

As long as the data marketplaces have privacy issues addressed, they will be viable and competitive even against the big players.

Various business models should be available regarding ownership of data and pricing. Various models such as subscriptions or pay per use. The pricing, the fees etc is a problem.

Due to the heterogeneity of the actors involved in a data marketplace (big variety of interested parties), it will be difficult to define the business models and have the right contracts for each one.

Ownership of data is also an issue. Especially when it comes to data that are produced by applying functions or processing over a dataset. If for example there is a service that performs dimensionality reduction, and someone applies it over a dataset giving specific parameters, how owns the produced dataset? How we deal with these or even more complex cases (combination of more than 1 dataset)?

Federated marketplaces is difficult to define and handle. There is always the competition factor. Probably a hierarchical model – the "bigger" marketplaces on the top and then smaller marketplaces under it, with some charging schema. Technically will not be a problem, there can be various solutions, but business-wise it remains a difficult challenges