



**HAL**  
open science

# Data Ecosystem for Industrial Product- Service Systems (IPS2) Based on a Decentralized Data Architecture

Lara Pörtner, Robert Möske, Andreas Riel

## ► To cite this version:

Lara Pörtner, Robert Möske, Andreas Riel. Data Ecosystem for Industrial Product- Service Systems (IPS2) Based on a Decentralized Data Architecture. *Procedia CIRP*, 119, pp.1228-1233, 2023, 10.1016/j.procir.2023.02.190 . hal-04162650

**HAL Id: hal-04162650**

**<https://hal.science/hal-04162650>**

Submitted on 15 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

33<sup>rd</sup> CIRP Design Conference

# Data Ecosystem for Industrial Product- Service Systems (IPS<sup>2</sup>) Based on a Decentralized Data Architecture

Lara Pörtner<sup>a,b,\*</sup>, Robert Möske<sup>b</sup>, Andreas Riel<sup>a</sup>

<sup>a</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 46 Avenue Félix Viallet, 38000 Grenoble, France

<sup>b</sup>CAMELOT Management Consultants AG, Gabrielenstraße 9, 80636 Munich, Germany

\* Corresponding author. Tel.: +49-1732698104; E-mail address: [lara.portner@grenoble-inp.fr](mailto:lara.portner@grenoble-inp.fr)

## Abstract

Industrial Product-Service Systems (IPS<sup>2</sup>) confront design organizations with new challenges. One of those is the need for co-designing products with services, and their underlying service provider and supplier networks including the business models driving their cooperation for value delivery to the end users. In this context, data collected about all the organizations involved in IPS<sup>2</sup> design and delivery are becoming increasingly valuable assets for both, the design as well as the continuous improvement, adaptation, and optimization of IPS<sup>2</sup>. In today's company IT infrastructures, data is typically managed in centralized data warehouses or data lakes. In order to cope with the dynamic nature of IPS<sup>2</sup> design and operation, agile de-centralized data organizations would be appropriate. Qualitative statements about the feasibility of some decentralized architectures realized with data mesh principles are given from practical use cases. This work focuses on a decentralized data architecture to enabling IPS<sup>2</sup> design and operation. Therefore, this work investigates the related challenges from centralized databases and proposes a decentralized concept solution tackling the requirements data privacy, data traceability, data accountability and data usability for the use case of a Mobility as a Service (MaaS) data ecosystem. For this purpose, compared to existing use cases, modern technology solutions such as blockchain technology and smart contracts are integrated in the concept.

© 2023 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer review under the responsibility of the scientific committee of the 33rd CIRP Design Conference

*Keywords:* Industrial Product-Service Systems; Decentralized Data Architecture; Data Ecosystem; Data-Enabled Design;

## 1. Introduction

Intelligent solutions driven by digitalization have a huge potential to foster e.g. climate protection targets [1]. Sharing services such as car sharing, bike sharing, or e-scooter sharing, as well as new mobility services such as flexible on-demand ride pooling services, play an important role for the transportation sector [2], which accounts for the second-largest emissions in EU [3]. These concepts act and store their data for other providers independently and separately. IPS<sup>2</sup> represent a new solution-oriented approach for delivering value in use to the customer during the whole life cycle of a product. They aim to combine value-adding services and products [4]. In this context, Mobility as a Service (MaaS), is not a completely new thought pattern that provides a new direction for personal

mobility improvements although, it can be seen more as an evolutionary continuation of the ongoing digital integration of information relevant for travelers [5]. Such MaaS solutions deliver users' transport needs through a single interface of a service provider [6, 7]

However, it is questionable how the increasing volumes of data in a data economy [8] can be sustainably collected, networked, and analyzed so that several providers can benefit from data sets and mobility becomes a service for the customer. According to a German study from 2022 regarding data economy [9], Business to Business (B2B) data sharing entails several obstacles. For example, in addition to incompatible datasets, lack of matching partners, or legal uncertainties have been seen as the biggest obstacles, while data privacy is also a

huge challenge when it comes to the exchange of data with others.

Data is often stored in centralized data bases, in Big Data Warehouses or Data Lakes [10]. To make organizations truly data oriented, a paradigm shift can be realized with the decentralized data mesh concept. Here, data is organized in domains and self-managing data teams which carry out their own work in an agile and product-oriented way [11].

This work therefore attempts to provide a decentralized design proposal for a data architecture based on existing fundamental design patterns that meets the requirements of the customer and the service provider in a data ecosystem (data privacy, data traceability, data accountability, data usability). For this purpose, principles of the data mesh concept are used and combined with blockchain technology and smart contracts as modern solutions to fulfil the requirements.

To provide such a concept, this paper proceeds as follows. Chapter 2 defines the terms Industrial Product-Service Systems, Data Mesh and Digital Ledger Technology. It then provides a quick overview of the current research status as well as the research question. Chapter 3 discusses the challenges of centralized data architecture and presents practical use cases of decentralized data architecture concepts. In chapter 4, the presentation of the derived decentralized data architecture concept is given. In chapter 5, an evaluation and discussion of the findings, limitations and an outlook on possible future research follows. Chapter 6 concludes the paper.

## 2. Background, Research Question and Methodology

### 2.1 Industrial Product-Service Systems

Industrial Product Service Systems (IPS<sup>2</sup>), specified as integrated product and service components, represent a new solution-oriented approach to delivering value to customers throughout a product's lifecycle [12]. They are characterized by the integrated and mutual planning, development, creation and use of product and service parts including software components in business-to-business applications. IPS<sup>2</sup> foster Circular Economy, leading to an increased lifetime, and more effective use of products and material. This typically goes with a break in ownership, which requires having more and precise data about the RUL (remaining useful life) of products. For this, data need to be collected all through the products' life cycle, and made coherent and useful [13]. Brissaud et al. [13] consider data-enabled design more appropriate in the context of IPS<sup>2</sup>, since design shall be value-driven, and data are just a means to achieve that. In this context, service life cycle data enable feedback of real service use patterns to the design of the underlying products, processes, and infrastructure.

### 2.2 Data mesh

Data mesh can be seen as a cultural shift in how companies think about their data. Instead of data being a byproduct of a process, it becomes a product (Product Thinking) [14]. The data mesh concept is promoted as a solution to the challenges that centralized data structures bring, such as accessibility and

organization of data [10]. Hence, the term “paradigm shift” associated with data mesh emerges [11].

The purpose of a data mesh is to use its transformation to solve the loss of the nature of data itself, the high costs associated with managing monolithic architectures (e.g., Data Warehouses or Data Lakes) [15], and the significant pressure on data teams. To achieve this, data mesh should be based on four core principles (cf. Table 1): domain-oriented decentralized data ownership and architecture, data as a product, self-service data platform, and federated computational governance [14].

Table 1. Data mesh principles.

|  |  |
|--|--|
| Domain-oriented decentralized data ownership and architecture [14], [16] | <ul style="list-style-type: none"> <li>Decentralized team composed of domain representatives, a clear ownership and distributed responsibility for each data product</li> <li>Organization's teams that can better understand and produce the data of their specific business domain</li> <li>Ownership is decentralized, and the provision of analytical data must always be coordinated with established domains</li> <li>Domains can both store and provide their data</li> </ul> |
| Data as a product [10], [14]   | <ul style="list-style-type: none"> <li>Concept of "Product Thinking" becomes the organization's top priority</li> <li>Analytical data is now seen as a product, and consumers of this data are now seen as customers to fulfill their needs</li> <li>Existence of a high-level infrastructure allowing the teams to produce and maintain their data products</li> </ul>  |
| Self-serve data platform [14], [17]                                      | <ul style="list-style-type: none"> <li>A platform is needed to independently manage the domains</li> <li>The self-service platform must be able to provide the tools and interfaces needed to create and maintain data products</li> </ul>   |
| Federated computational governance [17]                                  | <ul style="list-style-type: none"> <li>A governance model must be created which enables interoperability between different domains</li> <li>Execution of decisions and the decentralization and independence of each domain</li> <li>The federated computational governance concept is a complex model that does not reject change</li> </ul>  |

### 2.3 Distributed Ledger Technology

Distributed Ledger Technology (DLT) is a technology with a decentralized database. It allows transactions to be stored in a distributed decentralized system across all participants [18]. To guarantee data safety, each node checks a new transaction on validity [19]. This leads to greater transparency and thus to the reduction of information asymmetries within a supply chain [19, 20].

DLT is often used synonymously with blockchain, but blockchain is only the most common form of DLT concepts [21]. DLT's can be public or private; in a public DLT, a new node can join the system directly without being accepted by others. There are also DLT systems that are only accessible to a certain group of participants [22]. Moreover, DLT allows the application of smart contracts [21], i.e., digital programs that execute themselves when the terms of the agreement are met [23]. With this possibility, verification of compliance by transaction partners is guaranteed. In addition, smart contracts support integrity and trust across all participants as a result [19].

### 2.4 Research Question and Methodology

Research to date tends to focus on private blockchain solutions and has not yet conclusively clarified the use of public DLT and its technical specifications [24]. This again highlights the need for a data architecture that combines existing use cases and places them in the context of public DLT systems [25]. Therefore, this work considers a decentralized architecture with a public DLT system to fulfill the requirements of all supply chain participants by answering the following research question: *What is a possible design of a decentralized data*

architecture built on a public DLT system including data mesh principles to fulfil the requirements of participants of a data economy such as data security, data traceability, data accountability, and data usability for IPS<sup>2</sup>?

To answer this research question, functional requirements for the data architecture are first derived from literature and from the perspective of stakeholders of the data ecosystem of IPS<sup>2</sup>. Subsequently, an investigation of existing, implemented data architectures against the requirements takes place. With this knowledge, a proposal of a design concept that does not have the limitations of existing solutions and which is based on validated technological building blocks, can be achieved. Based on this, a critical evaluation of the concept against the requirements follows. The validation regarding the practical implementation still must be carried out.

### 3. Design of Decentralized Data Architecture Concept

#### 3.1 Challenges and requirements of data architectures

Many organizations manage their data in centralized databases (e.g., Big Data Warehouses, Data Lakes), which suffer from the problem of an overload of data teams in response to the growing needs of the organization. Even though data lakes are currently the most common type of data architecture used by organizations [14], they do not meet the latter's needs in terms of data ownership and traceability. More generally, in today's supply chains, there are two common ways of operating information systems:

- 1) Each supply chain participant manages their own data, without access or sharing with others.
- 2) Parties in the supply chain depend on a trusted third party to exchange data [21].

Both solutions are fraught with problems: the central entity must be trusted not to modify the data and to manage it securely [26]. As for management and liability, there is the question of responsibility in case of errors.

In case the central storage device fails, it affects the entire network, and a single point of failure appears [10, 27]. This leads to an unsatisfactory alignment between the architecture and the organizational needs. Therefore, the problems of the **loss of the nature of the data** itself, **high costs** related to the management of monolithic architectures, significant **pressure on data teams**, among others must be tackled [10]. Looking closer into practical use case like car sharing, requirements of participants of a supply chain have also been managed by building up a data architecture. There exists a huge conflict between a customer's privacy requirement and the demand of accountability. That means, customers would like to have data privacy assured while using a service, but service providers need to have the ability to know the real identity of a customer. In addition to **data privacy** and **accountability**, **traceability and usability** are also significant for the service provider side [24].

Therefore, the following functional (FR) non-functional (NFR) requirements can be derived for the design of a decentralized data architecture including data mesh principles (cf. Table 2).

Table 2. Requirements of data economy design concept.

|  |  |
|--|--|
| <b>FR 1: Data privacy</b>                | The design should preserve customer's data privacy. Service providers must only be allowed to see certain customer data and is not allowed to access all supply chain data.  |
| <b>FR 2: Data traceability</b>           | Stored transactions must not be modifiable and changes to the customer data must be traceable for other parties of the data ecosystem. Service providers must be able to trace both customer and product related data. |
| <b>FR 3: Data accountability</b>         | Customer data must be available with the transaction of the service.   |
| <b>FR 4: Data usability</b>              | Front-end access for customers and service providers must be given. Back-end access for service providers should be provided in addition.  |
| <b>NFR 1: Decentralized architecture</b> | The design should be realized with a decentralized database.   |

#### 3.2 Practical decentral architecture solutions

To build up a design for an architecture tackling all the challenges, it is important to understand in a first step how the requirements are already managed in already existing data architectures of organizations. In the following, four use cases are presented in which decentralized architectures are used. In each of the use cases, solutions to the challenges of the central architectures and the mentioned requirements can be found.

As a leading fashion platform in Europe, *Zalando* must store, process, and use significant amounts of data per day in the e-commerce business. For this purpose, *Zalando* has been using a data lake as central database for a long time. However, to counteract the disadvantages such as the lack of ownership of the data, the poor quality of the data after its processing and the organizational one, *Zalando* switched to its own data mesh approach. *Zalando* has implemented the treatment of data as a product and the provision of resources for quality assurance of the data with additional understanding of its use through some changes. The original central data lake storage has been retained and a metadata layer and governance containing information about it has been added. Thus, retaining the original centralized processing platform does not provide an all-encompassing decentralized architecture design. Users can also integrate their data with their data into the common infrastructure through Simple Storage Service (S3) buckets (FR 4). Information on specific frontend applications is not provided. Furthermore, clusters (spark clusters) are available on the processing platform, and users use these technologies without the team responsible for the infrastructure having to know what the users are doing (FR 1, FR 3). With regard to FR 2, *Zalando* creates a data infrastructure as a platform with an interoperable service environment [10,28].

*Netflix* as a streaming service provider must be able to handle and integrate large amounts of data. The user can use technologies such as GraphQL and Apache Iceberg to gain access to a metadata catalog from which it is possible to create his pipelines and access to a list of pipelining process standards (FR 3). Changes in the database trigger the entire process. Because these entities are owned by the customer, these entities transfer changes to the data repository and persist changes in an *Apache Iceberg table*. This provides an approach to ensure data privacy (FR 1). In order for the customer to find an easy handling of the service (FR 4), *Netflix* uses a topology with audit mechanism that is present in the data source, in the output of the processor and in the data batches stored in *Apache Iceberg format*. In this way, the accuracy is checked in the data source (FR 2)[10, 29].

*DAPA* is designed as a decentralized, accountable, and privacy-preserving architecture for car sharing services [24].

DAPA consists of five major phases: system setup, customer registration, car rental, car audit, and customer revocation. In the concept architecture, decentralized and dynamic validation servers are employed to assist in managing customers’ real identities instead of a single trusted authority to reduce the risk of the single point of failure and to build decentralized trust for customers (FR 1, NFR 1). Moreover, a new privacy-preserving identity management scheme (PPIM) achieves privacy preservation for customers (FR 1) and accountability for car sharing service providers (FR 3). The service provider has the possibility to verify the validity of customers’ ‘identifications without revealing customers’ real identities (FR 1). In addition, the concept is designed for efficiency in terms of computational costs and communication overheads. Even if the concept already includes data privacy elements, solutions such as smart contracts and blockchain but have not yet included it in their concept.

*Gaia-X* is an European initiative aligned with the EU Commission to provide an open, transparent and secure digital ecosystem, where data and services can be made available, collated and shared in an environment of trust [30]. Companies and citizens will collate and share data while keeping control over them with the implementation of secure federated identity and trust mechanisms (FR 1, FR 2). The architecture of *Gaia-X* is based on the principle of decentralization (NFR 1). A networked system linking many cloud services providers together is based on the values openness, transparency, and trust (FR 3). Therefore, aligning network and interconnection providers, Cloud Solution Providers (CSP), High Performance Computing (HPC) as well as sector specific clouds and edge systems is necessary. Easy access to the available providers, nodes and services is guaranteed and data will be provided through federated catalogues. Moreover, sovereign data services to ensure the identity of source and receiver of data make data accountability safe (FR 1, FR 3).

*Catena-X* is a project providing the first open and collaborative data ecosystem for the automotive industry of the future based on a *Gaia-X* compliant IDS-system [31]. It includes a global data space hosting industry, linking suppliers globally through an interoperable and trusted data-driven value chain. Here, the principle of data as a product is followed. The claim is data sovereignty and control over the data (FR 1, FR 2). As a central communication component of the *Catena-X* architecture, the Eclipse Data Space Connector (*EDC*) creates the connection for sovereign, cross-organizational data exchange, thereby implementing data ownership (FR 1). The network foundation is built upon shared services and *GAIA-X* / *IDS*A principles. *Gaia-X* and *Catena-X* show the first practical use case example for a general data economy system.

All these practical examples show the feasibility and added value of decentralized organization of data. **Error! Reference source not found.** evaluates requirements fulfilment (“X” means full fulfilment, “(X)” stands for partial fulfilment). It becomes visible that none of these decentralized concepts combines data mesh principles and the requirements of participants of an *IPS*<sup>2</sup> data economy in one architecture design.

Table 3. Requirements fulfilment of different practical use cases.

| Use case   | Data privacy | Data traceability | Data accountability | Data usability | Decentralized architecture |
|--|--------------|-------------------|---------------------|----------------|----------------------------|
| DAPA   | X            | (X)               | X                   | X              | X                          |
| Decentralized, accountable and privacy-preserving architecture   |              |                   |                     |                |                            |
| Zalando  | X            | X                 | X                   | X              | (X)                        |
| Maintain centralized data lake storage with addition of metadata layer and governance; including data mesh principles                          |              |                   |                     |                |                            |
| Netflix  | X            | X                 | X                   | X              | (X)                        |
| Prioritize reducing operational complexity rather than achieving best results in terms of cost and performance; including data mesh principles |              |                   |                     |                |                            |
| Gaia-X   | X            | X                 | X                   | (X)            | X                          |
| Open, transparent and secure digital ecosystem based on the principle of decentralization  |              |                   |                     |                |                            |
| Catena-X   | X            | X                 | (X)                 | (X)            | (X)                        |
| First open and collaborative data ecosystem with a decentralized governance across few organizations   |              |                   |                     |                |                            |

#### 4. Decentralized Data Architecture Concept for a Data Economy

In our design proposal, the fulfillment of the previously derived requirements is of key importance. Based on the practical use cases, it is possible to develop a data architecture with the help of the solutions prevailing in the literature (Data mesh, blockchain, smart contracts) to fulfill the given requirements for the design of a data economy system (cf. Figure 1).

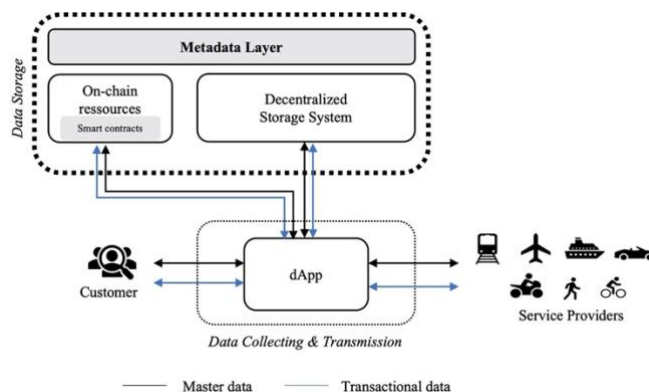


Figure 1. Decentralized data architecture design concept

##### 4.1 Data Collecting and Transmission

To ensure a decentralized distribution (NFR 1), most of the data is not collected via S3 buckets unlike the Zalando solution but stored directly in a decentralized storage system. The master data of the customer (e.g., name, address, address) must first be included in the application example of this work. An application example in the context of *MaaS* is the case that person A wants to travel from Munich to Hamburg as fast and climate friendly as possible. The platform should therefore propose the best possible service provider based on the master data provided. Thus, the way must be found via an application to first record master and transaction data of customers and service providers (e.g., airlines, car sharing providers, rail service providers, etc.).

All data is stored on the storage component so that all changes can be tracked. For ease of use for the customer, *dApp* can be used as an interface [32]. This frontend application

offers a simple solution for manually entering master and transaction data (FR 4) [33].

DApps have the advantage of being able to regulate access and thus provide each entity with individual user interfaces to ensure that not every service provider receives data indiscriminately. In addition, all affected entities are notified via the access portal as soon as a manual entry has been made. This input can then be verified and confirmed by the other entities. If an entity is still missing, the system automatically checks the entity and reminds it to complete the missing data (FR 2, FR 3). Consequently, this application also serves as a data provider for the entire supply chain, as it ultimately provides relevant data for confirmation by its controls [33].

#### 4.2 Data Storage

The resulting master and transaction data of the customer and the service provider are passed directly to the storage components, where they are stored and processed. The concept design includes modern technologies such as DLT technology and smart contracts as a solution proposal for a sustainable data ecosystem to master the following challenges: (1) Processing large amounts of data with high performance and low latency; (2) Scaling problems of a blockchain; (3) Storage of different types of data.

The use of a combination of off-chain and an on-chain component should provide a solution to process large amounts of data from the data ecosystem [27]. This intermediate component also solves the scaling problem of the blockchain [32]. In addition, it is also possible to store files of different types on the off-chain component (not only transaction data) [33]. This also allows, for example, the uploading of pictures of driver's license for the use of car sharing, which individual customers must also upload to the platform for control.

In order to be able to use a secure and immutable DLT at the same time, the off-chain storage structure is connected to a DLT, on which, however, only the hash values of the data and smart contracts with the corresponding input data are stored [28] (FR1, FR 2). Here, FR 1 is promoted by the modern solution with smart contracts. This also ensures the implementation of access restrictions, as not all entities in the supply chain are allowed to view all data. Smart contracts also regulate the exchange of data between the on-chain and off-chain components [34]. This data exchange can be made possible by oracles [35]. Data generated by data storage on the off-chain component follows the same principle as with a DLT. Thus, the off-chain generated hash values can be stored on the DLT [36].

The on-chain solution can be implemented via the Ethereum platform [27, 34]. This platform also enables the use of smart contracts. In addition, Ethereum supports various programming languages for smart contracts (e.g., Solidity), which is particularly advantageous in a system with many entities, as the entities can be used with preferred programming languages [34, 37]. An additional advantage is the support of various interfaces for communication between Ethereum nodes and the dApp. Consensus mechanisms are used to validate and confirm the data on the blockchain. The automatically and manually generated data stored via IPFS is validated using a special

consensus mechanism based on different trust centers [21]. Since all service providers want to avoid loss of reputation, they are encouraged to report the manually collected data correctly.

To describe the data of the on-chain resources, the data storage and the smart contracts, a meta data layer is used. This makes it possible, for example, to determine general data about new blocks in the chain, number of transactions or information about the integrity of the blockchain.

### 5. Discussion and Evaluation

All practical examples analyzed in chapter 3 show the feasibility and added value of decentralized organization of data. The design presented in chapter 4 incorporates elements of best practice examples and combines data mesh principles with modern technologies such as blockchain and smart contracts to organize data sets sustainably for the use case example MaaS in a data economy.

Table 4. Implementation of requirements in the design concept.

| <b>Data Mesh principles</b><br><i>(Data ownership, Data as a Product, Self-serve data platform, Federated computational governance)</i> |   |
|---|---|
| <b>FR 1: Data privacy</b>   | <ul style="list-style-type: none"> <li>Smart contracts regulate data exchange between on-chain and off-chain components.</li> <li>This also ensures the implementation of access restrictions, since not all entities in the supply chain are allowed to view all data.</li> <li>DApps have the advantage of being able to regulate access and thus also provide each entity with individual user interfaces to ensure that not every service provider receives data indiscriminately.</li> </ul> |
| <b>FR 2: Data traceability</b>  | <ul style="list-style-type: none"> <li>DApps as application for data recording.</li> <li>All entities concerned are notified via the access portal as soon as a manual entry has been made.</li> <li>Checking and confirmation of the input from the other entities possible.</li> <li>In case of missing input, the system automatically checks the entity and reminds them to complete the missing data.</li> <li>Meta data layer for data visualization of general process data.</li> </ul>    |
| <b>FR 3: Data accountability</b>  | <ul style="list-style-type: none"> <li>DApps have the advantage of being able to regulate access and thus also provide each entity with individual user interfaces.</li> <li>In case of missing input, the system automatically checks the entity and reminds it to complete the missing data.</li> </ul>   |
| <b>FR 4: Data usability</b>   | <ul style="list-style-type: none"> <li>For ease of use for the participants, dApp is used as an interface.</li> <li>IFRS and Ethereum serve as back-end solutions (DLT: on chain, off-chain solution).</li> </ul>   |
| <b>NFR 1: Decentralized architecture</b>  | <ul style="list-style-type: none"> <li>Decentralized data storage with data ownership of each participant.</li> </ul>   |

The whole concept tries to achieve a design that includes the socio-technological data mesh principles (*data ownership, data as a product, self-serve data platform, federated computational governance*).

**Data ownership** is related to the organization of data. That means, a decentralized architecture needs a clear ownership and distributed responsibility for each data product. In the design concept, every service provider is responsible for their data and not any other service provider can access other data indiscriminately. Moreover, **data as a product** is realized through the fact that data in the data economy are seen as a top priority product now, while consumers of these data as customers. The goal is to fulfill the customers' needs. Therefore, the concept design tries to fulfill all requirements of the data ecosystem. The principle of **self-serve data platform** requires a platform to provide the tools and interfaces needed to create and maintain data products. This platform is given through the whole design concept with its backend and frontend solution proposals.

Since Catena-X can already be seen as an approach to provide an open and collaborative data ecosystem, with **federated governance**, this work achieves an approach which



enables interoperability between different participants. This includes the execution of decisions and the decentralization and independence of each participant in the data ecosystem.

The proposed design concept attempts to map central requirements in IPS<sup>2</sup> and can thus be transferred to further application examples in the context of Industry 4.0 and flexible network production. Nevertheless, this work also has some limitations: First, its practical validation is only in a very early phase, since its deployment in practice requires adequate partnerships and IT infrastructure. Second, the processing of large amounts of data is of great importance in this data ecosystem and is realized via an on-off-chain solution. This work focuses on a public, permissionless solution, which is logic-oriented, while private DLT solutions have not been considered so far. Third, the scalability of the data ecosystem is based on a theoretical consideration only at this stage. Forth, this work did not yet investigate the skills required by designers to propose and implement the proposed concept in the context of IPS<sup>2</sup> design projects.

## 6. Conclusion and Outlook

We have proposed a decentralized data architecture that provides the enabling basis for modern IPS<sup>2</sup> with MaaS as an exemplary use case. The essential unique contribution lies in covering the requirements of a data ecosystem i.e., data privacy, data traceability, data accountability, and data usability. In our future work, we aim at validating our proposed concept with pilot partners from automotive industry in the context of Industry 4.0. From an education perspective, we will investigate how designers can be trained and empowered to be able to consider the proposed solution concept.

## References

- [1] Fauser, J., N. Sigle, and D. Hertweck, "DATA-BASED APPLICATION SCENARIOS FOR E-SCOOTERS", ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences VIII-4/W1-2021, 2021, pp. 41–47.
- [2] Bauer, G., C. Zheng, J.B. Greenblatt, S. Shaheen, and D.M. Kammen, "On-Demand Automotive Fleet Electrification Can Catalyze Global Transportation Decarbonization and Smart Urban Mobility", *Environmental Science & Technology* 54(12), 2020, pp. 7027–7033.
- [3] "Topic: Transportation emissions in the EU", Statista. <https://www.statista.com/topics/7968/transportation-emissions-in-the-eu/>
- [4] Smith, G., J. Sochor, and I.C.M. Karlsson, "Adopting Mobility-as-a-Service: An empirical analysis of end-users' experiences", *Travel Behaviour and Society* 28, 2022, pp. 237–248.
- [5] Smith, G., "Making Mobility-as-a-Service: Towards Governance Principles and Pathways", 2020. <https://research.chalmers.se/en/publication/516812>
- [6] "[PDF] Sampo Hietanen. CEO, ITS-Finland - Free Download PDF", <https://silo.tips/download/sampo-hietanen-ceo-its-finland>
- [7] Jittrapirom, P., V. Caiati, A.-M. Feneri, S. Ebrahimigharehbaghi, M.J.A. González, and J. Narayan, "Mobility as a Service: A Critical Review of Definitions, Assessments of Schemes, and Key Challenges", *Urban Planning* 2(2), 2017, pp. 13–25.
- [8] Oliveira, M.I.S., and B.F. Lóscio, "What is a data ecosystem?", *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, ACM (2018), 1–9.
- [9] Dr. Rohleder, B., "Bitkom-Charts\_Datenökonomie\_04\_05\_2022\_final.pdf", 2022. [https://www.bitkom.org/sites/main/files/2022-05/Bitkom-Charts\\_Datenökonomie\\_04\\_05\\_2022\\_final.pdf](https://www.bitkom.org/sites/main/files/2022-05/Bitkom-Charts_Datenökonomie_04_05_2022_final.pdf)
- [10] Machado, I.A., C. Costa, and M.Y. Santos, "Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures", *Procedia Computer Science* 196, 2022, pp. 263–271.
- [11] Machado, I., C. Costa, and M.Y. Santos, "Data-Driven Information Systems: The Data Mesh Paradigm Shift", pp. 6.
- [12] Meier, H., R. Roy, and G. Seliger, "Industrial Product-Service Systems—IPS 2", *CIRP Annals* 59(2), 2010, pp. 607–627.
- [13] Brissaud, D., T. Sakao, A. Riel, and J.A. Erkoyuncu, "Designing value-driven solutions: The evolution of industrial product-service systems", *CIRP Annals* 71(2), 2022, pp. 553–575.
- [14] "How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh", [martinfowler.com. https://martinfowler.com/articles/data-monolith-to-mesh.html](https://martinfowler.com/articles/data-monolith-to-mesh.html)
- [15] "Data lake governance: A big data do or die", *SearchCIO*. <https://www.techtarget.com/searchcio/feature/Data-lake-governance-A-big-data-do-or-die>
- [16] Moses, B., "What is a Data Mesh — and How Not to Mesh it Up", *Medium*, 2021. <https://towardsdatascience.com/what-is-a-data-mesh-and-how-not-to-mesh-it-up-210710bb41e0>
- [17] "Data Mesh Principles and Logical Architecture", [martinfowler.com. https://martinfowler.com/articles/data-mesh-principles.html](https://martinfowler.com/articles/data-mesh-principles.html)
- [18] "The limits of trust-free systems: A literature review on blockchain technology and trust in the sharing economy", *Daneshyari*. <https://daneshyari.com/en/article/6854021>
- [19] Roeck, D., H. Sternberg, and E. Hofmann, "Distributed ledger technology in supply chains: a transaction cost perspective", *International Journal of Production Research* 58(7), 2020, pp. 2124–2141.
- [20] Kannengießer, N., S. Lins, T. Dehling, and A. Sunyaev, "Trade-offs between Distributed Ledger Technology Characteristics", *ACM Computing Surveys* 53(2), 2021, pp. 1–37.
- [21] Sunyaev, A., *Internet Computing: Principles of Distributed Systems and Emerging Internet-Based Technologies*, Springer International Publishing, Cham, 2020.
- [22] Corsten, H., and S. Roth, *Handbuch Digitalisierung*, Vahlen, 2022.
- [23] Szabo, N., "Formalizing and Securing Relationships on Public Networks", *First Monday* 2(9), 1997.
- [24] Huang, C., R. Lu, J. Ni, and X. Shen, "DAPA: A Decentralized, Accountable, and Privacy-Preserving Architecture for Car Sharing Services", *IEEE Transactions on Vehicular Technology* 69(5), 2020, pp. 4869–4882.
- [25] Kannengießer, N., M. Pfister, M. Greulich, S. Lins, and A. Sunyaev, "Bridges Between Islands": Cross-Chain Technology for Distributed Ledger Technology", pp. 10.
- [26] Vergne, J., "Decentralized vs. Distributed Organization: Blockchain, Machine Learning and the Future of the Digital Platform", *Organization Theory* 1(4), 2020, pp. 263178772097705.
- [27] Fridgen, D.G., D.N. Guggenberger, D.T. Hoeren, et al., "OPPORTUNITIES AND CHALLENGES OF DLT (BLOCKCHAIN) IN MOBILITY AND LOGISTICS", pp. 244.
- [28] "Data Mesh in Practice: How Europe's Leading Online Platform for Fashion Goes Beyond the Data Lake", *Databricks*. [https://www.databricks.com/session\\_na20/data-mesh-in-practice-how-europes-leading-online-platform-for-fashion-goes-beyond-the-data-lake](https://www.databricks.com/session_na20/data-mesh-in-practice-how-europes-leading-online-platform-for-fashion-goes-beyond-the-data-lake)
- [29] Netflix Data Mesh: Composable Data Processing - Justin Cunningham,
- [30] "Gaia-X", <https://www.data-infrastructure.eu/GAIA/Navigation/EN/Home/home.html>
- [31] "Home | Catena-X", <https://catena-x.net/de/>
- [32] Chauhan, A., O.P. Malviya, M. Verma, and T.S. Mor, "Blockchain and Scalability", 2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C), IEEE (2018), 122–128.
- [33] Hepp, T., P. Wortner, A. Schönhals, and B. Gipp, "Securing Physical Assets on the Blockchain: Linking a novel Object Identification Concept with Distributed Ledgers", *Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems*, ACM (2018), 60–65.
- [34] Augusto, L., R. Costa, J. Ferreira, and R. Jardim-Goncalves, "An Application of Ethereum smart contracts and IoT to logistics", 2019 International Young Engineers Forum (YEF-ECE), IEEE (2019), 1–7.
- [35] Kannengießer, N., S. Lins, C. Sander, K. Winter, H. Frey, and A. Sunyaev, "Challenges and Common Solutions in Smart Contract Development", *IEEE Transactions on Software Engineering* 48(11), 2022, pp. 4291–4318.
- [36] Nair, A.R., R. Gupta, and S. Tanwar, "FAIR: A Blockchain-based Vaccine Distribution Scheme for Pandemics", 2021 IEEE Globecom Workshops (GC Wkshps), IEEE (2021), 1–6.
- [37] Westerkamp, M., F. Victor, and A. Kupper, "Blockchain-Based Supply Chain Traceability: Token Recipes Model Manufacturing Processes", 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), IEEE (2018), 1595–1602.