Energy Informatics

METHODOLOGY

Open Access

Exploring decentralized data management: a case study of changing energy suppliers in Germany

Linda Rülicke[1*], Florian Fehrle[1], Arne Martin[4], Antonello Monti[2,3], Volker Berkhout[1], Oliver Warweg[4] and Sven Möller[4]

*Correspondence:
linda.ruelicke@iee.fraunhofer.de

[1] Energy Informatics, Fraunhofer IEE, Joseph-Beuys-Straße 8, Kassel 34117, Hessen, Germany
[2] Institute for Automation of Complex Power Systems, RWTH Aachen University, Mathieustraße 10, Aachen 52074, NRW, Germany
[3] Center for Digital Energy, Fraunhofer FIT, Mathieustraße 10, Aachen 52074, NRW, Germany
[4] Energy Informatics, Fraunhofer IOSB-AST, Am Vogelherd 90, Ilmenau 98693, Thüringen, Germany

Abstract

This paper presents an innovative approach to decentralized data management in the German energy market, focusing on the use of decentralized data management with the help of Data Spaces to facilitate the automated change of energy suppliers within 24 h. The central focus of this research is the MakoMaker Space, a demonstrator project that employs the Connector from the Eclipse Data Space Components. The MakoMaker project demonstrates the successful automation of energy supplier changes, emphasizing the preservation of customer data sovereignty. It shows an alternative approach to the process, putting the customer into the center. Customers retain control of their data, which is accessible to providers as needed. While the paper discusses the potential for further enhancements, such as the integration of an identity provider and the development of a sustainable business model for service coordination, the primary focus is on the demonstrator's successful application in a pilot setting.

Keywords: Decentralized data management, Data space, Market communication, Energy supplier change, Energy sector

Introduction

Motivation and problem description

In the ever-evolving digital data universe we are witnessing, the increasing demand for innovative, efficient solutions to get the full potential of the available data, distributed among multiple actors. In this context, data ecosystems, aiming at facilitating collaborative data sharing, have emerged. To succeed in this idea, an efficient orchestration of decentralised stored data assets accompanied by a robust trust framework is necessary. This approach offers a possibility to share and manage data, while assuring that the data owner preserves their sovereignty over their data.

Additionally it contributes to European ideas. Europe's diverse energy market, characterized by its commitment to sustainability and cross-border energy exchange, necessitates robust, interoperable data ecosystems. This research, while rooted in the German energy sector, echoes the European Union's broader objectives of digital sovereignty and a single digital market. It underscores the potential of decentralized data management to enhance energy market efficiency, security, and consumer empowerment across Europe.

This research paper presents a methodology to use decentralised data management with the help of an enabling technology called Data Space, on the exemplary use case to change an energy supplier in the German energy domain within 24 h.

Data Space technology means a significant paradigm shift in how we perceive, manage, and harness data. It promotes the development of an integrative ecosystem that brings together disparate data sources into a singular, easily accessible, and secure Data Space, whilst ensuring that the data provider maintains control over data usage. This technology tackles the limitations associated with traditional data silos and cultivates a more collaborative, interoperable, and trust-oriented data-sharing environment, emphasizing transparency and control.

This paper gives an outline of a promising use case to show the advantages of a Data Space in the energy sector. We take a deeper look into market communication processes and picked the change of an energy supplier as an example. Ultimately, our aim is to take the theoretical concept and go the next steps towards a real life implementation.

Paper structure

The paper gives an overview of data ecosystems and their current relevance in the energy sector. The background chapter establishes the theoretical foundation necessary for understanding the broader concept. It begins with a definition of data ecosystems. This is followed by an exploration of Data Spaces, the technical concept to implement a data ecosystem. The sections "Data spaces in the energy sector" builds the bridge to sector specific viewpoints from the energy sector and how Data Spaces and decentralised data management can contribute to those domain-specific questions. The next section explores the current market communication framework, delving into the specific terms and conditions pertinent to the German energy market. The background chapter concludes with a synthesis of related work and identifies a research gap, underscoring the importance of applying Data Space concepts in practical, real-world contexts.

Chapter "Methodology" presents the methodology used to implement an example of decentralised data management in the energy sector. This chapter outlines the necessary assumptions and prerequisites for the case, detailing the roles of the involved actors and the data they manage. In the chapter "Results" the technical details of the implemented use case are described such as the software components, the process for switching energy contracts, and the mechanics of decentralized data management. The final two chapters discuss the results and their implications, summarizing the main insights and suggesting perspectives for future research.

Background

Data ecosystems

The idea of data ecosystems brings the concept of business ecosystems to a next level. The term of business ecosystems was framed by Moore in 1993 (Moore 1993). He states that different actors in a ecosystem influence each behaviour and evolution.

If we take for example the relationship between a cheetah and a gazelle, we see that their speed is influenced by each presence in the same habitat. By catching slow gazelles the cheetah is increasing the gene pool of fast attributes in the specie of gazelles. As a consequence the species of cheetahs need to become faster in order to not extinct (Ebner 2006).

We see the interaction of the actors in a more collaborative way, than in the previous given example. Nevertheless to adapt the idea to a business context, technology is needed to enable the communication between the participating actors. Isherwood and Coetzee framed the term "Digital Business Ecosystem" (Isherwood et al. 2011). It is "a decentralised environment where very small enterprises and small to medium sized enterprises interoperate by establishing collaborations with each other". Collaborations play a major role in the development of Digital Business Ecosystems where it is often difficult to select partners, as they are most likely strangers. Even though trust forms the basis for collaboration decisions, trust and reputation information may not be available for each participant (Isherwood et al. 2011).

The overarching goal is to create a collaborative environment for business participants, where they can confidently engage with one another without needing to fully know their counterparts. This level of collaboration and trust is facilitated by a robust technological framework. We conclude that Digital Business and data are intricately linked, to the extent that effective collaboration is only possible with the ability to exchange data. Therefore the term "data ecosystems" presents the final definition we will give and use throughout the remaining paper. Oliveira et al. (2018) conducted a meta-analysis of the term data ecosystem and define it as follows:

> *a set of networks composed by autonomous actors that directly or indirectly consume, produce or provide data and other related resources (e.g., software, services and infrastructure). Each actor performs one or more roles and is connected to other actors through relationships, in such a way that actors collaboration and competition promotes data ecosystem self-regulation.*

The term was firstly used by Vision (2010) in 2010, closely followed by Ding (2011) in 2011. The later have created a Semantic Web-based platform for supporting linked open government data Linked Open Government Data (LOGD). This portal serves as an open-source infrastructure for LOGD production and consumption, and as a community portal for a global network of developers, data curators, and users. The paper introduces the portal and highlights its innovative features and the lessons learned from its development.

Data spaces

Data spaces can be understood as an enabler or tool to achieve the goal of a vivid data ecosystem.

Nagel and Lycklama (2021) define a Data Space "as a decentralized infrastructure for trustworthy data sharing and exchange in data ecosystems based on commonly agreed principles." Compared to the definition of a data ecosystem, Data Spaces emphasize the infrastructure component. To effectively collaborate "commonly agreed principles" are defined and agreed upon while participating in a specific Data Space. Otto adds an essential characteristic, by stating, that the data will not be fully integrated, accepting possible data redundancies (Otto et al. 2022, p.7).

Apart from the fundamental component of data, a data ecosystem comprises many diverse actors, each playing a distinctive role within the structure. The Reference Architectur of the International Data Space Association (IDSA) structures those actors into four categories: (1) Core Participant, (2) Intermediary, (3) Software Developer, and (4) Governance Body (IDSA 2022)

Core Participants are divided into two primary roles: the Data Supplier and the Data Customer. The Data Supplier acts as a Data Creator, Owner, and/or Provider, responsible for generating and making data available and thereby providing the raw material to fuel the ecosystem. In contrast, Data Users or Customers, utilize the provided data to extract valuable insights, make informed decisions, or develop data-driven products and services.

A second category is the Intermediary role, which can be understood as an enabler or in a broader view as a platform for other participants. This role is exclusive to trusted organizations that foster mutual benefits for all participants. They achieve this by building trust, providing metadata, and developing business models.

The third category is filled by technological Service Providers that contribute software to the Data Space. These companies take on business roles such as App Developers and Connector Developers, providing essential technical support and innovation.

Finally, the last category is the Governance Body. This entity is responsible for establishing decision making processes and enforce rules and guidelines within the Data Space. Their role is crucial in creating a trust framework that underpins the sustainable operation of the Data Space, ensuring that data exchange is both secure and efficient (IDSA 2022)

The design principles of a Data Space are anchored in several fundamental concepts, each playing an important role in shaping its architecture and functionality (Nagel and Lycklama 2021).

Firstly, the principle of Data Sovereignty is fundamental and central. This concept revolves around the idea that individuals or entities have a right to control their data. In practical terms, this means that within Data Spaces, participants have the power to decide what data they want to share, with whom and under what terms. They keep control over their data, ensuring that they can make it accessible to others under their own terms. This principle is vital in maintaining trust and integrity within the data ecosystem.

Another critical principle is Data Sharing and Exchange on equal terms. This concept seeks to level the playing field in terms of data access and utilization. The goal is to create an environment where the value and competition are driven by the quality and relevance of data and services, rather than by monopolistic control over vast data reserves. Such an approach encourages a cooperative working environment, where competition is healthy and based on innovation and quality. This principle is crucial in preventing the formation of data silos.

Lastly, the principle of a decentralised infrastructure is integral to the vision of Data Spaces. Unlike traditional centralized systems, the envisaged infrastructure is a decentralized network of interoperable systems. These systems are bound by a common set of functional, technical, operational, and legal standards, enhancing the system's overall efficiency and integrity. This decentralized approach not only facilitates greater flexibility and scalability but also enhances security and privacy measures.

Technically, this infrastructure appears as a network of API-based IT platforms. Within these platforms, users can manage data flows and harness the economic potential of the data they control (Nagel and Lycklama 2021).

Together, these principles contribute to a secure, equitable, and efficient digital ecosystem where data is not just an asset but a tool for innovation, collaboration, and growth (Nagel and Lycklama 2021).

As the definitions of data ecosystems and Data Spaces are very broad and sector agnostic a taxonomy was developed by Gelhaar et al. (2021) in order to better classify and focus the work in each Data Space. It consists of seven key dimensions and eighteen characteristics and can be seen in Table 1.

Data spaces in the energy sector

As the energy sector evolves towards more decentralized structures, Data Spaces emerge as an fitting technological solution due to their decentralized nature as well. This alignment allows for an effective mapping of physical infrastructures to their digital counterparts, ensuring a more integrated and efficient system. Moreover, with the rise in energy trading between different countries within Europe, there is a growing interest from the European Union in fostering seamless data exchange of energy information across national borders. This interest is driven by the need to optimize and streamline energy distribution and consumption across the continent, making Data Spaces an important component in the evolving landscape of European energy management and collaboration.

The creation of a Common European Energy Data Space (CEEDS) is one of the priorities set by the European Commission in the Digitalization of Energy Action Plan European Comission (2022). As also shown by the use case covered in this paper, a Data Space approach is a key element to facilitate the empowerment of citizens in the energy sector. This is true for two fundamental reasons:

1. Data Spaces facilitate the data sovereignty at customer level, thanks to their distributed approach
2. Data Spaces are also the best vehicles to go across different domains enabling cross-domain use cases

Table 1 Data Space Taxonomy visualized as a morphological box Gelhaar etal. (2021)

Meta-dimension	Dimension	Characteristics			
Economic	Domain	Scientific	Government	Industry	
	Purpose	Innovation	Interaction	Transaction	
	Organization	Keystone	Platform	Marketplace	Decentralized
Technical	Infrastructure	Centralized	Distributed		
	Openness	Open	Closed		
Governance	Inter-dependence	Tightly Coupled	Loosely Coupled		
	Control	Central	Decentral		

The success of CATENA-X (Ganser and Göller 2023) and the Mobility Data Space (Lemm 2023), along with the European Union's action plan calling for a closer connection between the energy and automotive sectors, positions Data Spaces as a promising technology for facilitating this cross-domain integration.

Market communication framework

Since a complete overview of the framework of the communication in the German electricity market would exceed the capacity of this paper we will only discuss the relevant parts, regarding the supplier change. This paper uses the roles of the German electricity market as described in the role-model of the market communication (der Energie- und Wasserwirtschaft 2023) to describe the process. The regulatory framework of the German electricity market is formalized in several laws and acts. The most relevant law for this use case is the energy industry act [(dt. Energiewirtschaftsgesetz (EnWG)]. The paragraphs §§6-10 EnWG describe the unbundling process and separate the distribution system operator (Distribution System Operator (DSO), dt. "Netzbetreiber"), the electricity supplier (ES, dt. "Lieferant") and the Metering Point Operator (MPO) (dt. Messstellenbetreiber) from each other. The reason for this unbundling process as defined in §6 EnWG is that the grid represents a natural monopoly, which could be abused if the Electricity Supplier (ES) and the DSO are the same entity. Since the MPO has data from all ES in a grid the same reasoning as previous holds true. As the DSO might gain an advantage by providing wrong metering data to the ES, the role of the MPO was carved out.

The data exchange in the German electricity market is currently mainly done with standardized data exchange protocols and the data format Electronic Data Interchange for Administration, Commerce, and Transport (EDIFACT). The current data exchange is defined in the Market Communication 2022 ("MaKo 2022") from the German Federal Network Agency. Bundesnetzagentur (2023)

Process for electricity supplier change

The process of ES change in Germany is a communication intensive process, as it includes five different parties, who generate and exchange data. The active participants in this process are the electricity end customer, the new ES (Electricity Supplier New (ESN)), the old ES (Electricity Supplier Old (ESO)), the DSO and the MPO. The process is documented within the of processes to deliver electricity to the customer ("Geschäftsprozesse zur Kundenbelieferung mit Elektrizität" (GPKE)), which are provided by the German Federal Network Agency ("Bundesnetzagentur").Bundesnetzagentur (2023) The current process flow is summarized in Fig. 1.

The process is initiated as the customer orders a new contract from an ESN. The ESN then terminates the old contract at the ESO. The ESO confirms this termination and provides the date of termination. Then the ESO deregisters the Market Location (MaLo) by informing the DSO about the end of supply. If the deregistration was successful, the DSO confirms the end of the contract. This process is done simultaneous to the registration of the new supplier, which if successful is confirmed by the DSO. In this case the new supplier can send a contract confirmation to end customer. The DSO provides the MPO with the information about the change of supplier for the given MaLo. In return the MPO transmits the data of the energy values at the time of change to the DSO if the

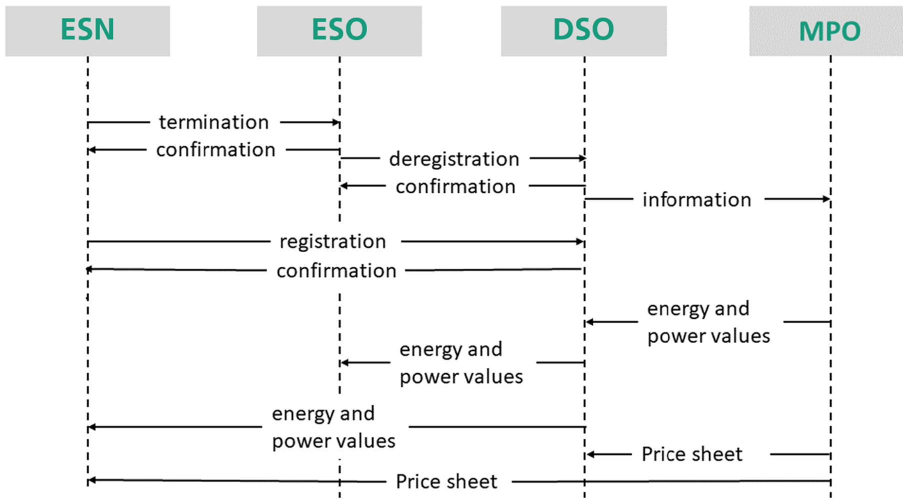

Fig. 1 Current supplier change process (Bundesnetzagentur 2023)

supplier change is done during a year. These values are then sent from the DSO to the ES. The ESO then uses this data to generate the final bill for the end customer, while the ESN has a start point for the following billing processes. At the end of the process the ESO sends a final invoice to the end customer (Bundesnetzagentur 2023).

Shortcomings in the current process

The current supplier change process has two main shortcomings. The first one is that the process is rather slow and can currently extend to a maximum of three weeks as defined in §20a EnWG. The main reasons for this long period of time are the communication of the participants with each other and that each participant has to verify the identity of the end customer. With the new regulation "Lieferantenwechsel 24 h", which is also defined in §20a EnWG, the German Federal Network Agency wants to address the issue by requiring the process to be done in 24 h and on every workday. The readiness of the current processes, methods and technologies to comply with this is currently debated (Bundesnetzagentur 2023). The second main issue is that the personal data such as the name and address of a person that should belong to the end customer is transmitted between the participants without direct knowledge or possibility to intervene of the end customer. This does not align with the goal to make the use of personal data transparent as defined in the general data protection regulation under Art.5 1a (Parliament 2016).

Related work

In this paper, we endeavor to bridge the gap between the conceptual groundwork detailed above and the practical realities of implementing Data Spaces in the field. Our research predominantly concentrates on examining various methodologies aimed at achieving decentralized data management within the energy sector, leveraging a range of technological frameworks. A notable reference architecture is that proposed by the IDSA, which advocates for the integration of all endpoints through the use of an IDSA compatible connector, thereby enabling each participant to function within the data space. The Reference Architecture Model, as delineated by the IDSA, is characterized by the absence of a central authority overseeing data management (IDSA 2022). This architectural paradigm distinctly contrasts with centralized data management systems, such as data lakes, and decentralized data networks that operate without unified governance or rules. By exploring these alternative methodologies, we aim to get insights into the critical success factors.

Scheibmayer and Deindl (2010) called the concept they developed "Internet of Energy", which aims at a decentralized IT-infrastructure in order to establish an efficient inter-company data exchange. In their research they found, that a major inefficiency in the German energy market is the lack of standardisation regarding the communication and storage of data. Their proposed "Internet of Energy" aims to fix that problem by adding three helping systems. The Energy Interface System (EIS) provides standardised interfaces for the participants to set up on their end. The Energy Name Service (ENS) provides a role model, that maps each participant with a role, providing the functionality of quickly determining with who you communicate. The third component introduced in the system is the Energy Security Service (ESS), which facilitates secure communication among participants. By integrating this along with the other two added systems, the approach proposes a process that significantly reduces manual intervention and standardizes contract changes. The example chosen to demonstrate this enhanced process is an ES change, showcasing how these components work together to streamline and secure the procedure.

Kim et al. (2010) presented a secure decentralized data-centric information infrastructure for smart grids. The work explores the transition to a publisher-subscriber system within decentralized data-centric infrastructure, enhancing data delivery efficiency and scalability by using peer-to-peer networks. It relies on publishers announcing data availability and subscribers expressing interest, with matches facilitated by an information middleware. This middleware also introduces a cost-efficient, scalable, distributed storage solution utilizing numerous smaller-capacity physical disks. This approach supports a distributed storage network across various grid entities, each with unique data storage requirements. Control within this network is maintained through a structured peer-to-peer grid overlay of trusted nodes, ensuring secure and reliable data delivery. This approach is not considering the customer but mainly the companies currently involved in the communication of the energy sector.

A study from Aslam et al. (2021) on data management in the context of local energy communities. The authors introduce a decentralized data management framework, integrating technologies like blockchain, Distributed Hash Table (DHT), role-based access control, and encryption. It focuses on securely storing and managing data, with encrypted data on the Distributed Hash Table (DHT) and metadata on the blockchain. This framework supports multi-level data access and ensures real-time data security. The paper also includes a security and performance analysis, demonstrating the framework's effectiveness and scalability.

Pemsel (2019) analysed how the blockchain technology can improve the concept of electricity supplier change. He found that the blockchain framework can be used for the process as it is especially useful in the verification of the end customer and secure against manipulation attempts. In his work Pemsel mentioned that the DSO would be the authority in the consensus method, since this role is regulated and closely supervised by the state. Non-compliant ES would be excluded from the blockchain infrastructure as a way to discourage the behaviour. Issues that were not tackled in the thesis are the topics of how the MaLo would be hidden from market participants who do not have a legitimate interest, how to include the MPO in the process and if a sustainable business case can be created for the application of the blockchain technology in the process.

Feser (2023) conducted a comprehensive evaluation of the EDC-Connector in his bachelor thesis, specifically within the context of the energy sector. His primary focus was on evaluating the performance, specifically regarding the transmission speed of big data time series. His findings reveal that certain data planes provided by the EDC are more suitable for the given task than others.

Conclusion and research gap

The research presented in this paper demonstrates the use of various technologies to establish decentralized data management within the energy sector.

Scheibmayer and Deindl (2010) also explored the concept of changing energy suppliers but adopted a less decentralized approach. They centralized data storage and allowed customer access through a web interface, with customer requests subsequently forwarded to relevant actors.

Aslam et al. (2021) concentrated on small energy communities. They limited participant numbers to maintain reasonable calculation times within their blockchain architecture, a crucial factor for the efficiency of their system.

Pemsel (2019) took a different route by excluding personal data from the equation. Instead, he opted to record all data openly on the blockchain. However, this method does not align with the data sovereignty approach that our research aims to achieve, as it potentially compromises individual data control and privacy.

Lastly, Feser's work Feser (2023) delved into a more technical analysis, assessing the advantages and limitations of the EDC-Connector. This study provided valuable insights into the technical aspects and potential of the EDC-Connector within the context of decentralized data management in the energy sector.

In concluding the related work section, it becomes evident that our research aims to navigate through the challenges presented by decentralized data management. Drawing upon the insights from previous studies, we aim to bridge the theoretical and practical aspects of this field. Our focus on utilizing Data Space technology, particularly within the energy sector, is a response to the identified gaps and challenges highlighted in these works. By synthesizing the knowledge gained from these studies and applying it to our research, we strive to transform Data Space technology from a largely theoretical model into a practical, implementable solution suitable for real-world pilot scenarios.

Methodology

The following section describe a scenario where a common German energy Data Space would be established. In the German energy market, customers looking to switch their energy supplier often face a lengthy and cumbersome process that can last up to a maximum of three weeks. This delay primarily stems from the complex interplay of various market actors, each with their own sets of data and verification protocols.

The planned regulation of “Lieferantenwechsel 24 h” mandates that this transition must be completed within a 24-h window. To comply with this directive, a radical change of existing systems and procedures is required, emphasizing technological enhancements and automation across different players in the energy sector.

A major bottleneck in the current process is the verification of customer data. This step is time-consuming, largely due to the need to reconcile disparate versions of customer information held by various entities in the energy market. Often, there is a significant challenge in ensuring that all parties refer to the same customer.

Use case description

A customer can use the Data Space to change the energy supplier quickly. The data of the customer would not be stored centrally on a platform, but the single-source-of-truth of this data set would stay within the infrastructure of the customer or an infrastructure that is controlled by or on behalf of the customer. Later within the chapter we explain, why we included the customer into the process and discuss advantages and disadvantages of this decision. To facilitate this scenario, we are utilizing Data Space technology. The proposed use case envisions a data exchange where the data stays with the data owner and therefore provides a single source of truth of the often discussed data points.

Use case conditions

Assumptions The use case was designed under the following assumptions:

- A general German energy Data Space is available
- Each legal entity participating in the process has its own Data Space connector
- Master data is kept with the Data Owner
- Entities which are using data, for which they are not the Data Owner, can load local copies of the data
- The coordination of the energy supplier change is done by a service provider

In light of our assumptions, we identify key participants in the use case as follows: the customer, MPO, DSO, ESN, and ESO. This configuration marks a significant departure from the current process outlined in the earlier chapter, primarily by integrating the customer into the process. This integration is crucial, as the customer is positioned as the only legitimate entity to act as the Data Owner for customer data. This change represents a major shift in the German market communication, where the customer has traditionally been excluded.

Our objective is to establish a ’single source of truth’, where this truth is maintained by the data owner, underscored by the principle that the owner is responsible for their own data. This approach prompts an essential question: in scenarios where the customer is absent from the process, who then assumes ownership of customer data? To address this, we propose integrating the customer directly into the process, thereby allowing for the decentralised storage of customer data at the customer’s end. This strategy not only empowers customers but also enhances data integrity and ownership, aligning with our goal of achieving a more transparent and efficient data management system in the energy sector.

Actors

Each actor which is involved in the process, would hold different data. A comprehensive view of the corresponding data categories is listed in Table 2. In addition to the energy-specific roles within the system, there's the inclusion of a Service Provider, who is coordinating the change process. This role, essential for the smooth transition and implementation of processes within the system, can be executed by various companies. It's important to note that while this role is significant within the context of the example provided, it is not deemed a central role for the entire system. This distinction underscores the system's design, which is modular and allows for different entities to assume this role as needed.

Mapping the taxonomy

After the previous chapter gave a clear picture of the business process that is to be achieved, a classification of the taxonomy shown in Table 1 can be conducted. This helps us to adapt our application design to the requirements of the use case. Analytical applications have, for example, different requirements towards atomic operations than transactional systems. Therefore our analysis is structured according to the three meta-dimensions: Economic, Technical, and Governance.

Economic Under the Economic dimension, the domain is identified as Industry, given that the primary data exchange occurs between various companies and customers. Although the prototype was developed in a research context, the final deployment of the use case falls outside of the scientific environment. The purpose is transactional, characterized by the exchange of standardized messages among participants. The organization of interactions is decentralized, with no central actor coordinating all exchanges.

Technical In the Technical dimension, we refer to the foundational principles of the technical infrastructure supporting data exchange. The infrastructure is distributed, with each participant contributing to the process. Data, some of which is personal and thus necessitates careful handling, is exchanged on a need-to-know basis. Therefore, the level of openness is classified as closed due to these restrictions on data accessibility.

Governance Under the Governance dimension, the actors within the proposed data ecosystem demonstrate loose coupling. The control dimension, which pertains to authority over critical data resources, is classified as Decentral due to the dispersed nature of the data resources within this use case.

Table 2 Types of data for each data space participant

Data space participant	Data
Customer	Name data
	Address data
Energy supplier old	Contract data
Energy supplier new	Contract data
Meter operator	Meter master data
	Metered Data
System operator	Market location data
Service provider	Mapping table to route the data

An overview of the final classification can be found in Table 3.

Results

The following chapter presents the implementation of the described use case within a Data Space. The project is referred to as MakoMaker.

Connector technology

From a technical perspective, the immediate decision to make is selecting the appropriate connector software to implement the described functionality. In this project, the EDC-Connector, developed by the Eclipse Foundation, was used for four main reasons:

- The EDC-Connector is an open-source project, which is constantly updated to add new and improved functionality to the connector.
- It provides a very customizable framework, to which you can easily add your own so-called "extensions".
- It uses a very precise path with regards to policies and data access.
- The IDSA rates it in their Data Connector Report with a maturity level of Technology Readiness Level (TRL) 8-9 (Giussani and Steinbuss 2023).

This leads to very clear benefits of using this connector as a starting point for the demonstrator. The EDC-Connector provides several sample projects to get better insights into the features of the connector. In this project we used one of the provided connectors with a http-push communication (Eclipse 2023) to act as the basic connector, on top of which we added extensions to add the desired functionality for the MakoMaker Space.

The EDC-Connector does provide two options for data transfer between connectors: The first is the 'provider push' method, where the provider pushes the data via a HTTP POST request to a provided data server which then processes it. The second method is 'consumer pull', where the provider opens a proxy to the data and the customer retrieves the data via a HTTP GET request with an authentication code received from the data provider.

We opted for the 'provider push' method for the sake of simplicity, since it requires a few less steps on the customer side to get the data and gives us the opportunity to create a separate data sink to which the data can be pushed and processed therefrom. To grasp the functioning of the software, it is essential to comprehend how the EDC-Connector manages stored data: the initial step to publish data involves creating an asset. Assets

Table 3 Classifying the energy supplier use case in the data space taxonomy

Meta dimension	Dimension	Characteristic
Economic	Domain	Industry
	Purpose	Transaction
	Organization	Decentralized
Technical	Infrastructure	Distributed
	Openness	Closed
Governance	Interdependence	Loosely coupled
	Control	Decentral

are the EDC-Connectors way of storing data. They either hold data themself or hold an access link to the data. In our case, it holds a URL to a Data Source Server running in the background of the connector. The asset itself cannot be accessed by another connector. In order to be accessible, two more elements are needed: a policy definition and a contract definition. The policy definition states under what conditions the asset can be accessed. For example, in our case we use policies to restrict access to the data to certain connectors. The contract definition connects the asset with the policy definition, so that the connector can map all elements together.

If all three are created, another connector can access the catalog of the connector, in which all available assets are visible with their corresponding policies.

The MaKo Maker Space project, as previously described, utilizes the EDC-Connector as its foundational technology. We have leveraged the EDC-Connector's extension framework to integrate the specific functionalities required for the MaKo Maker Space. Upon initialization, the EDC-Connector activates multiple ports, including a management port for controlling the connector, a protocol port for inter-connector communication, and a versatile API-Port, which is central to our discussion.

The architecture of each connector in the demonstrator is divided into two primary components. The 'connector side' manages core data-related tasks such as policy enforcement, secure data exchange, and data discoverability through a local catalog. Conversely, the 'API side,' accessible via the API-Port, automates the MaKo Maker Space's operations. It employs JSON messages and HTTP POST requests for inter-party communication, enabling the execution of various tasks by the connector side, such as data access or publication.

Presently, two key functionalities are embedded in the Demonstrator: the creation of new energy contracts and the termination of existing ones. Both processes extensively utilize the API side of the connector, as depicted in Figs. 3 and 4. Additionally, the system enables the current supplier to retrieve up-to-date personal customer data, a feature detailed in the chapter on decentralized data management.

Changing an energy contract

The main functionality of the MakoMaker Space is to provide a simpler, more automated way for customers to change an energy contract to a new energy supplier. The Fig. 2 describes the adapted high-level process of changing an energy contract. It was derived from Fig. 1, whilst adapting the processes of a data space and focusing the process on the consumer.

To automate this process, it is necessary to break the process down into separate parts and provide lower-level instructions for implementing the connector software. Figure 2 already does the first part by separating the process of contract termination (green) and making a new contract (orange) into two parts. Figures 3 and 4 are depicting each step in more detail focusing on the main actors.

Figure 3 shows the low-level flow of a new contract being created. For that to be happening, the customer must make the assets with his personal information accessible to the supplier via a contract definition. The contract definition references a policy, that gives reading rights to the new supplier only. Therefore the data can only be accessed by the ESN. The supplier can then negotiate a contract for this data. The ESN from then on

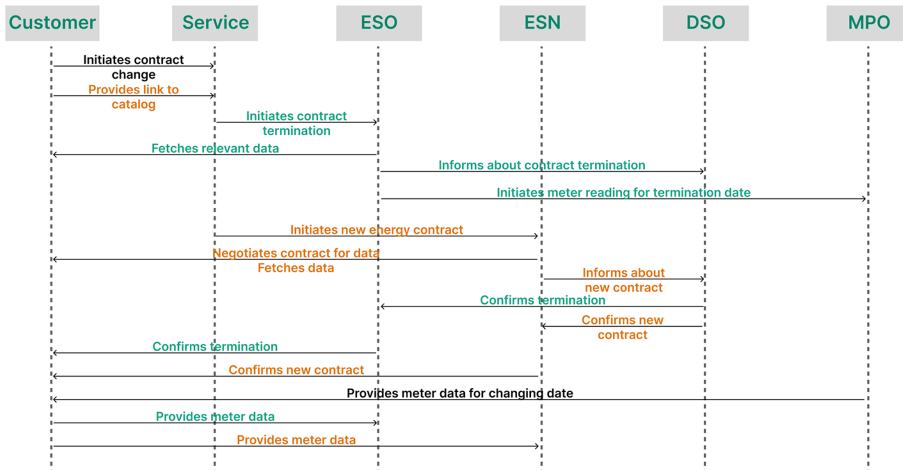

Fig. 2 Process overview energy supplier change

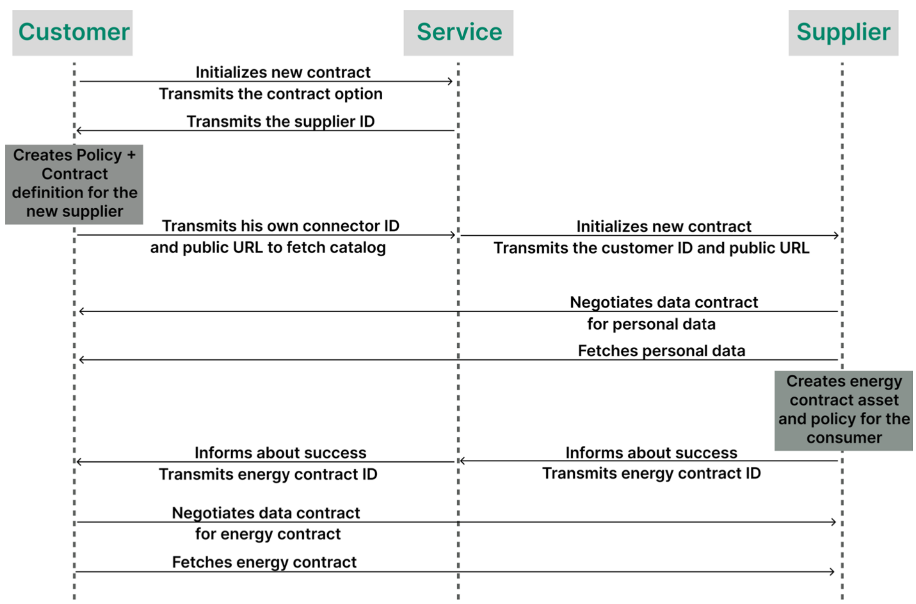

Fig. 3 Making a new energy contract

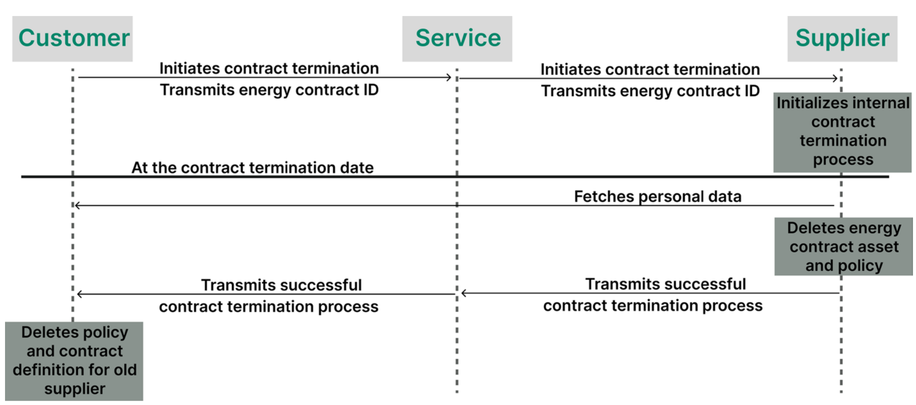

Fig. 4 Terminating an existing energy contract

can request the data, whenever it is needed, e.g. whilst creating bills. On his side he creates an energy contract, to which he gives access to the customer in form of an asset in his catalog. The customer can then negotiate a contract to access this file and from then on can get the data for the energy contract.

Figure 4 illustrates a similar process for terminating an energy contract, where the contract definition is deleted rather than created. Notably, this action does not instantly revoke access rights to the data. Access remains as originally specified in the initial contract agreement, a crucial aspect for both parties to meet legal data storage requirements. To prevent indefinite data access, a default time limit, such as 10 years, could be implemented. This would effectively restrict long-term data access, ensuring compliance while maintaining necessary legal obligations.

Decentralized data management

One of the main focuses of the MakoMaker Space is, that the data owner is the one who holds the data and is the only one allowed to grant access to it. An alternative is giving data away to a central authority, which processes and stores the data. An example of this would be a data lake. A data lake is a centralized storage solution within an organization, designed to store a variety of data in its native format, with the intent of future value extraction. Much like data spaces, in a Data Lake, data need not be pre-integrated (Mathis 2017).

Data Lakes primarily focus on the central aggregation of an organization's data. Consequently, considerations like data sovereignty and cross-company usage agreements, which are pivotal in Data Spaces, are typically not central to Data Lake solutions. However, it is feasible that a Data Lake could interface with a data space, adapting to data space protocols to selectively share content. Mathis (2017) A third option is the current solution used within the energy sector. As described before, personal data from the customer (name and address data) as well as the meter data, to which the customer has the ownership rights, is interchanged between parties in the energy market, without the customer having the full power over the data. The MakoMaker Space tries to solve that, by bringing the data source for the personal data back to the customer. In the case of the demonstrator, this is achieved by storing the data on the customer side and then making it available through an asset, that can be fetched by the parties needing it (mainly the energy supplier). This ensures, that the master data, representing the single source of truth, is always on the customer side, but the energy supplier can access the data.

This makes the process for getting the current data from the customer automated and thus less complicated. In the system, as it is in operation now, if a customer would change his name, he would need to inform his energy supplier. This is nowadays mostly done via E-Mail. Nevertheless, the customer must write an E-Mail with his new name to the energy supplier, where an employee must process this E-Mail and update the data in the data storage of the supplier. This is not an extensive task, but on a scaled-up operation level, it requires a lot of human resources, that could be used more efficiently. The MakoMaker Space solves this problem: Since the personal data now is stored by the customer, the customer would only need to change the data at his end, for the data to be updated. When the supplier needs this data, he would simply fetch this from the customer and therefore would have current data, without needing to update data manually.

This decreases the workload for changing data on the customer and the provider side, which would be beneficial for both parties.

Discussion

This paper presents a comprehensive study of the MakoMaker Space, focusing on the implementation and functionality of connector software, specifically in the context of energy contract management and decentralized data management. The choice of the EDC-Connector was made due to its open-source nature, its customizable framework that allows for specific extensions, precise policy and data access control, and a high maturity rating from the IDSA. The concept of decentralized data management was put into practical application through a sample project utilizing http-push communication from the EDC-Connector. This project was subsequently tailored to meet the unique requirements of the MakoMaker Space, effectively demonstrating the viability and functionality of the concept in a real-world setting. The paper also outlines the automation process for changing energy contracts, breaking it down into two key phases: the termination of an existing contract and the creation of a new one, demonstrating the practical application of the connector software in implementing these processes.

A big advantage of the proposed approach is the establishment of a single version of the truth. This benefits mainly energy suppliers, network operators and metering point operators as it is reducing the administration effort to handle a supplier change. In the proposed model, the customer initiates the change of energy supplier, either directly via the new energy supplier or through a third-party service that offers this feature. The use of a Data Space in the case of a third-party service provider presents a significant advantage: the customer only needs to share the link to their data. This does not require the service provider to access the data directly; they merely receive and forward the link to the relevant parties.

Upon receiving the link, the ESO can access necessary data given their rights to do so. This supplier then notifies the system operator about the termination of delivery and requests meter data from the meter operator for the termination date. Concurrently, the service provider facilitates contact between the customer and the ESN. Both parties negotiate a contract regarding data usage and acquire the requisite data to legally formalize the agreement.

An additional benefit of the described method is the precise identification of the customer through the connector address. This approach renders the time-consuming task of harmonizing customer data, as mentioned earlier, unnecessary. By utilizing the connector address, all participants within the Data Space can easily identify the correct customer through a clearly defined attribute, streamlining the identification process significantly.

Additionally the approach enhances data sovereignty and security. By distributing data across multiple nodes, rather than centralizing it, the risk of large-scale data breaches is significantly mitigated, safeguarding consumer information. This increases the power of the customer by reacting to growing societal concerns about data privacy in an increasingly interconnected world.

Our research fits in the currently evolving domain of Data Spaces. The Data Space Radar published by the International Data Space Association lists 43 Data Spaces in

10 different sectors (https://internationaldataspaces.org/adopt/data-spaces-radar/.) and highlights the breadth of this development. Especially the energy sector is increasingly interfacing with other domains, thereby striving for cross-sector interoperability. The synergy between the energy and automotive sectors, as evidenced by initiatives like Catena-X (Ganser and Göller 2023) and the Mobility Data Space (https://mobility-dataspace.eu/.), is an example of this trend. Given the mandates of the EU Action Plan for closer collaboration, such alliances are not only beneficial but necessary.

During the implementation of the MakoMaker project we found some limitations, which offer insights into areas for future development and refinement.

The demonstrator's functionality, while operational, currently excludes comprehensive data flows for both the DSO and the MPO. This exclusion, primarily a strategic choice for simplification and to focus on the evolving role of the customer, results in the absence of crucial data such as metering information.

While the integration of the customer offers benefits, it also presents challenges, particularly in terms of the effort and cost involved. One option of this integration involves customers directly installing and maintaining the necessary IT-infrastructure, specifically an environment to run a Java application. In doing so, the customer would need to ensure some level of minimal availability. However, this approach is likely to be feasible for only a small percentage of users, given the technical skills and resources required. The prospect of widespread adoption appears limited due to these inherent demands. A more common scenario would likely involve a service provider hosting the data connector on behalf of the customer. In this setup, customers access their data through a web interface provided by the service provider. This method eases the technical burden on the individual customer but introduces additional costs. These costs, stemming from the services provided by the host, would likely be reflected in the customer's energy bill. Additionally an easy usability for the customer would be crucial for a widespread adoption.

An important consideration here is the cost-benefit equation. The viability of this system hinges on whether the additional costs incurred by using a service provider are offset by the savings from automating the energy supplier change process. If the additional costs are lower than the savings, the customer would still benefit financially from the system. However, if the costs outweigh the savings, this could pose a significant limitation to the system's attractiveness and feasibility from a customer standpoint. Thus, ensuring that the financial benefits surpass the additional expenses is crucial for the system's acceptance and success.

This limitation, while notable, provides valuable direction for further development. They highlight the need for ongoing enhancements to ensure that the system not only meets the current requirements but is also robust and secure for future applications.

Conclusion and outlook

The project successfully demonstrates the feasibility of fully automating the process of changing energy suppliers using decentralised data management. A key achievement of this implementation is the maintenance of customer data as the single source of truth, which remains with the customer and is accessed as needed by the data provider. This approach ensures that any updates made to a customer's personal data are immediately available to all participants, enhancing the efficiency and accuracy of the data exchange. Another advantage is clear identification of the customer by the connector address.

As we look forward to advancing this field of study, our research roadmap encompasses several key areas, each aimed at deepening our understanding and application of Data Spaces in the energy sector.

Decentralized Energy System. Our immediate focus will involve developing a specific use case centered around an energy process. We plan to equip a Smart Meter with a novel connector, enabling us to delve into the practicalities and benefits of machine communication for decentralized energy turbines. This will provide a tangible example of our concepts in action, bridging the gap between theory and practice.

Comparative Performance Analysis. To objectively assess the efficacy of our approach, we will engage in a comparative analysis using measurable metrics. Parameters like overall process time and message throughput will be critically examined against existing technical alternatives. This will provide a quantifiable basis for evaluating the advantages and potential limitations of our system.

Scalability Assessment. A crucial aspect of our future work will be exploring the scalability of our system, particularly when thousands of parties are interacting. This will include examining the system's behavior in the face of network interruptions and varying network latency configurations, providing insights into its robustness and efficiency at scale.

Comprehensive Security Analysis. Given the importance of cybersecurity, a detailed analysis of the security implications of our decentralized approach will be conducted. We aim to identify potential new attack vectors and assess whether a decentralized system contributes to overall system robustness.

Practical Application Considerations. Finally, we will address multiple practical aspects of our solution, including maintainability and the ease of updating source code. This will help in understanding the real-world applicability of our system, ensuring that our research is not only theoretically sound but also practically viable.

Regulatory and Policy Implications. Our roadmap includes evaluating existing policies for their adaptability to decentralized data management in the German energy market, alongside a security analysis to determine optimal processes for decentralization. This dual approach will identify which energy market processes can benefit from transitioning to a decentralized setup and which should remain centralized, ensuring compliance with regulatory standards and enhancing system security.

Business Model. A significant area that remains to be explored is the development of a viable business model for the service coordinating the supplier change. Future research should focus on evaluating potential business models for this actor, addressing the challenges of sustaining the service financially while maintaining its effectiveness and integrity. This exploration is crucial for the long-term success and scalability of the automated energy supplier change system.

Through this comprehensive research agenda, we aim to offer in-depth insights into the research area, thoroughly exploring its opportunities and challenges.

Another interesting path for future research is exploring the organisational and technological implications of managing a large number of connections, particularly in scenarios where each household provides data through its own connector. Future studies are planned to assess the organizational effort required, as well as the technological demands in terms of bandwidth and server capacity, to effectively handle such a widespread and decentralized system. This research will be crucial in understanding the scalability and feasibility of the proposed data management approach.

Abbreviations

IDSA	International data space association
IEC	International electronical commission
MPO	Metering point operator
ES	Electricity supplier
ESN	Electricity supplier new
ESO	Electricity supplier old
MaLo	Market location
DSO	Distribution system operator
EDC	Eclipse data space components
TRL	Technology readiness level
CEEDS	Common European energy data space
EDIFACT	Electronic data interchange for administration, commerce, and transport
EnWG	Energiewirtschaftsgesetz
DHT	Distributed hash table
LOGD	Linked open government data

Acknowledgements
For the creation of this paper Large Language Models were used as a supportive tool. The work performed by the models was: improving the wording and sentence structure

Author contributions
LiRu: Writing: Chapter 1; 2 excluding 2.4, 2.3; 4; 6; 7. FlFe: Writing: Chapter 3, 5, 6. ArMa: Writing: Chapter 2.4, 3. AnMo: Writing: Chapter 2.3, Review. VoBe: Conceptualization, Review. OlWa: Conceptualization, Review. SvMo: Resources, Review.

Funding
Open Access funding enabled and organized by Projekt DEAL. This work was supported by Tennet TSO GmbH, the Fraunhofer Cluster of Excellence Integrated Energy Systems (CINES) and the ENERSHARE project (Grant Agreement No 101069831).

Availability of data and materials
Not applicable.

Code availability
We plan to publish the project MakoMaker Space under an open source license on the Fraunhofer IEE Github repository in the future. (https://github.com/FraunhoferIEE).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Received: 28 December 2023 Accepted: 13 February 2024
Published online: 22 February 2024

References

Aslam S, Bukovszki V, Mrissa M (2021) Decentralized data management privacy-aware framework for positive energy districts. Energies 14:7018

Bundesnetzagentur. Geschäftsprozesse zur kundenbelieferung mit elektrizität (gpke) (11/22/2022–1/10/2023). https://data.bundesnetzagentur.de/Bundesnetzagentur/DE/Beschlusskammern/BK06/BK6_83_Zug_Mess/831_gpke/bk622128_gpke_ba.pdf

Bundesnetzagentur. Verfahren zur festlegung von regelungen für einen beschleunigten werktäglichen lieferantenwechsel in 24 stunden (lfw24): Bk6-22-024 (2023). https://www.bundesnetzagentur.de/DE/Beschlusskammern/BK06/BK6_83_Zug_Mess/8353_Lieferantenwechsel/BK6_Lieferantenwechsel24h_node.html

der Energie-und Wasserwirtschaft eV, Anwendungshilfe BB (03/06/2023) Rollenmodell für die marktkommunikation im deutschen energiemarkt: Arbeitsgrundlagen marktkommunikation. https://www.bdew.de/media/documents/2023-03-06-AWH-Rollenmodell_MaKo_V2.1_BcwsudV.pdf
Ding L et al (2011) Twc logd: a portal for linked open government data ecosystems. J Web Semantics 9:325–333
DRM Datenraum Mobilität. Mobility data space. https://mobility-dataspace.eu/
Ebner M (2006) Coevolution and the red queen effect shape virtual plants. Genet Program Evolvable Mach 7:103–123. https://doi.org/10.1007/s10710-006-7013-2
Eclipse EDC (2023) transfer-03-provider-push. https://github.com/eclipse-edc/Samples/tree/main/transfer/transfer-03-provider-push
European Commission. Digitalising the energy system—eu action plan (2022). https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13141-Digitalisierung-des-Energiesektors-EU-Aktionsplan_de
Feser J (2023) Prototypical evaluation of interoperable transmission of energy data using eclipse dataspace components | sap blogs. https://blogs.sap.com/2023/11/23/prototypical-evaluation-of-interoperable-transmission-of-energy-data-using-eclipse-dataspace-components/
Ganser O, Göller F (2023) Das ziel: Ein einheitlicher datenraum: Catena x. Frankfurter Allgemeine Zeitung. https://www.faz.net/pro/d-economy/d-mobility/das-ziel-ein-einheitlicher-datenraum-19407305.html
Gelhaar J, Groß T, Otto, B (2021) A taxonomy for data ecosystems. Hawaii International Conference on System Sciences 2021 (HICSS-54). https://aisel.aisnet.org/hicss-54/os/managing_ecosystems/2
Giussani G, Steinbuss S (Zenodo, 2023) *Data Connector Report*. https://zenodo.org/record/7825732
IDSA. Ids ram 4 (2022). https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0
International Data Space Association. Data spaces radar: Faster ids breakthroughs are within range. https://internationaldataspaces.org/adopt/data-spaces-radar/
Isherwood D, Coetzee M, Venter HS (ed.) (2011) Enhancing digital business ecosystem trust and reputation with centrality measures. In: Venter H S (ed.) 2011 Information Security for South Africa (ISSA 2011), 1–8 (IEEE, Piscataway, NJ)
Kim Y-J, Thottan M, Kolesnikov V, Lee W (2010) A secure decentralized data-centric information infrastructure for smart grid. IEEE Commun Mag 48:58–65
Lemm K (2023) Binnenmarkt für daten soll die verkehrsbranche besser vernetzen: Pausenlos produzieren menschen informationen, wenn sie unterwegs sind. besonders sammelfreudig zeigt sich das auto. Heise—MIT Technology Review. https://www.heise.de/hintergrund/Binnenmarkt-fuer-Daten-soll-die-Verkehrsbranche-besser-vernetzen-9201713.html
Mathis C (2017) Data lakes. Datenbank-Spektrum 17:289–293. https://doi.org/10.1007/s13222-017-0272-7
Moore JF (1993) Predators and prey: a new ecology of competition. Harvard business review 75–86. https://parsmodir.com/wp-content/uploads/2019/08/innecho-moore1993.pdf
Nagel L, Lycklama D (2021) Design principles for data spaces—position paper
Oliveira MIS, Lóscio BF, Janssen M (ed.) (2018) What is a data ecosystem? (ed. Janssen, M.) Proceedings of the 19th Annual International Conference on Digital Government Research Governance in the Data Age, ACM Other conferences, 1–9 (ACM, New York, NY)
Otto B, ten Hompel M, Wrobel S (2022) Designing Data Spaces. Springer International Publishing, Cham. https://library.oapen.org/handle/20.500.12657/57901
Parliament TE (04/27/2016) & the Council of the European Union. Regulation (eu) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679
Pemsel J (2019) Anwendung der blockchain-technologie im umfeld der energiewirtschaft am beispiel des stromlieferantenwechsels
Scheibmayer M, Deindl (eds) M (2010) An ICT architecture to support business processes in the Internet of Energy. https://ieeexplore.ieee.org/document/5756548
Vision T (2010) The dryad digital repository: published evolutionary data as part of the greater data ecosystem. Nat Proc https://www.nature.com/articles/npre.2010.4595.1

Publisher's Note