

Designing Data Trustees: A Prototype in the Building Sector

Michael Steinert¹^a, Anna Maria Schleimer^{1,2}^b, Marcel Altendeitering¹^c and David Hick³

¹Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany

²TU Dortmund University, Dortmund, Germany

³DKSR GmbH, Berlin, Germany

Keywords: Data Trustee, Data Intermediary, Data Sharing, Data Space, Data Ecosystem.

Abstract: There are still major concerns about the sharing and use of personal data, even though it has great value for society. This is particularly evident in the context of buildings, where data on citizens' energy consumption offers great potential for optimization and resource conservation. However, building owners are reluctant to share their data due to concerns about control or misuse. Unlike business relationships in data ecosystems, where companies can establish technological trust mechanisms such as authorization and policy management, individuals require other parties to do so. The EU Data Governance Act proposes the use of neutral intermediaries called data trustees. However, the concrete design of data trustees for personal data remains open. To address this, we propose a prototype based on design science research methodology and data space technologies. The prototype demonstrates a data trustee for trusted sharing and use of personal data, with the added capability of leveraging decentralized service providers to offer value-added services, such as the generation of energy certificates. These decentralized services extend the functionality of the data trustee by providing adaptable, advanced solutions that benefit multiple stakeholders. In addition, the study contributes requirements and lessons learned for future implementations.

1 INTRODUCTION


Personal data has significant potential to create value for individuals and society, but its use is often hindered by concerns about privacy, trust and legal obligations (e.g. GDPR). In contexts such as the ongoing European energy crisis, access to energy data can support citizens by enabling services that improve energy efficiency (European Commission, 2022). However, creating value from personal data is challenging due to individuals' reluctance to share data, fears of losing control, and a complex regulatory landscape. Beyond regulatory constraints, different stakeholders, such as data suppliers and technology providers, must interact securely and transparently to foster trust and shared value (Moore, 1993; Curry, 2016, p. 35).


The process of generating energy certificates for buildings exemplifies these challenges (Li et al., 2019). To produce these energy certificates, which are critical information sources for assessing energy use and guiding improvements, building owners must


share sensitive data with municipalities and other entities. Building owners may be concerned about revealing private details, worry about data misuse, or face technical hurdles in sharing information. Even organizations struggle with inefficient data sharing due to data silos, inconsistent formats, and poor interoperability (EnergyREV, 2020; Heuninckx et al., 2023). These complexities underscore the need for secure, sovereign, and easy-to-use solutions that respect privacy and build trust among all stakeholders.

This contribution addresses the research question: *How to design data trustees for personal data in the building sector.*

We focus on data trustees - neutral intermediaries who establish a fair balance between the interests of all parties involved and enable a trusted exchange of data, including the necessary access (Federal Ministry of Education and Research Germany, 2022). Using a Design Science Research (DSR) approach (Peppers et al., 2007), we develop and evaluate a data trustee prototype using decentralized service providers. The prototype aims to provide building owners with direct, policy-based control over their data usage, streamline data flows from municipali-

^a <https://orcid.org/0009-0008-3888-2092>

^b <https://orcid.org/0000-0002-3264-8034>

^c <https://orcid.org/0000-0003-1827-5312>

ties, and support value-added services (e.g., generating energy certificates). In doing so, it overcomes key limitations of current data sharing approaches, such as manual data sourcing and the lack of standardized mechanisms.

Unlike existing data sharing solutions, which often rely on building owners to manually collect and provide their data to service providers, our prototype automates and streamlines these processes by acting as a trusted intermediary (i.e., data trustee) between building owners, municipalities, and energy certificate issuers. First, building owners no longer need to search for and compile billing and consumption data themselves; instead, they can leverage their municipality's data directly. Second, by integrating a certified energy certificate service, our approach ensures the accuracy and credibility of issued energy certificates, eliminating the risk of errors associated with manual data entry. Third, our architecture supports broader ecosystem engagement, allowing building owners to easily and securely donate their building data for municipal planning or other beneficial community efforts. In sum, the prototype not only reduces complexity and effort for individual building owners, but also provides a sustainable and open infrastructure for expanding data use cases, ultimately overcoming the limitations and fragmentation often found in traditional data sharing approaches.

2 BACKGROUND

2.1 Data Ecosystems and Data Trustees

Data ecosystems describe "a set of networks composed of autonomous actors that directly or indirectly consume, produce, or provide data and other related resources" such as services (Oliveira and Lóscio, 2018, p. 4). A fundamental aspect is the loose coupling of members and interdependencies, which distinguishes an ecosystem from strictly defined, fixed relationships. Ecosystems have different relationships between the goals of their members, their interdependencies, and their overall network, which together affect aspects such as governance and control (Bogers et al., 2019). Typically, data value chains in ecosystems focus on business and societal purposes and can "enable collaboration among diverse, interconnected participants that depend on each other for mutual benefit." (Curry, 2020, p. 7). Data ecosystems build on the various interactions and data exchange relationships of data owners and consumers. To realize these relationships, a software technology foundation is required. Data spaces and the underlying data infras-

tructures provide such technological concepts, aiming at tools to balance the interests of data providers with the use of data for a greater common good (Otto and Burmann, 2021). In data ecosystems such as data spaces, data trustees are a relevant type of actor in the context of personal data. Data trustees, also known as data trusts, are considered as legal structures that provide "independent stewardship of data" (Hardinges et al., 2019). They mediate access to data according to "contractually agreed or legally prescribed data governance regulations [...]" (Blankertz and Specht-Riemenschneider, 2021).

In addition, data trustees can complement data spaces by facilitating trusted data exchange in business-to-consumer (B2C) environments, ensuring that individual customers and their interests are prioritised. While data spaces focus on enabling secure and sovereign data exchange between companies (B2B), data trustees extend this to trusted data exchange with individual customers.

Figure 1 visualizes the relationship between the data space and infrastructure, the data objects exchanged, and the ecosystem participants, including the data trustee. Within this context, four distinct archetypes of data trustees emerge: first, data broker trusts, which act as neutral intermediaries between data owners and data users, granting access to those with legitimate requests; second, data processing trusts, which ensure secure and compliant data processing, typically in a business-to-business context. Additionally, data aggregation trustees combine data from different owners, while data custody trustees protect sensitive personal data (Lauf et al., 2023, p. 10).

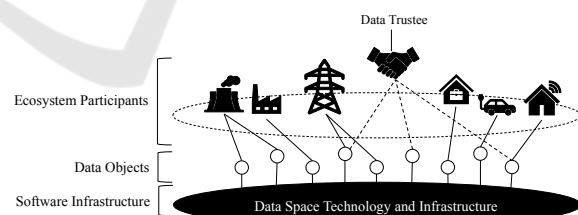


Figure 1: Data Trustees in the Context of Ecosystems and Data Spaces, own adaption based on (Otto and Burmann, 2021).

2.2 Urbanization & Sustainable Buildings

As global urbanization continues to increase, cities and urban areas play a critical role in creating sustainable and healthy environments and achieving the United Nations Sustainable Development Goals (Moran et al., 2018; United Nations, 2015). In ur-

ban areas, there is a constant demand for more housing as more people move to cities (Jenks and Jones, 2010). An uneven distribution of income and cultural activities between rural and urban areas drives this trend, leading to a demographic mismatch between the two (Porru et al., 2020). As a result, many cities have alarming carbon footprints and unsustainable operations (Moran et al., 2018). In addition to urban transportation, housing is a significant source of greenhouse gases in cities. The building sector is responsible for approximately 40% of total energy consumption (Li et al., 2019). Given this, housing is an important sector for improving the sustainability of urban areas and achieving global climate goals. One of the ways in which the European Union is trying to achieve this goal is through energy certification of buildings and "... achieving the great unrealized potential for energy savings in buildings" (European Union, 2010, p. 1). These energy certificates are based on several pieces of information, including a building's thermal characteristics, heating and air conditioning systems, ventilation, and indoor climate conditions (European Union, 2010). The data is then aggregated into a standardized energy certificate. Energy certificates provide several benefits to all stakeholders. They help buyers and renters assess the future energy costs of a building or apartment. For building owners, they provide a basis for making informed decisions about the need for renovations and the best ways to improve a home's carbon footprint. Municipalities and governments can use the certificates to set minimum energy performance requirements and set targets when buildings undergo major renovations (European Union, 2010). In addition, energy certificates can support urban planning and improve the energy consumption of specific areas (e.g. by offering subsidies for houses that suffer most from high energy consumption) (Jenks and Jones, 2010). Creating energy certificates requires collecting data from multiple sources, which can be challenging because data is often stored in silos, such as municipalities, energy providers, or building owners. This makes data difficult to access, especially with legal restrictions such as the Data Governance Act (DGA) and the GDPR, which limit access to sensitive data.

3 RESEARCH APPROACH

In this contribution, we present the prototype developed during three cycles of a DSR study following established guidelines (Peffer et al., 2007). DSR is a suitable approach for creating artifacts that address a "heretofore unsolved and important business prob-

lem" (Hevner et al., 2004, p. 84). It is therefore well suited for creating innovative solutions and prototypes in unexplored areas (Hevner et al., 2004). Data trustees represent such an unsolved business problem because they lack accessible design knowledge and guidance for their creation (Stachon et al., 2023). This lack of knowledge and concrete examples for creating data trustees using decentralized service providers serves as the problem-centric entry point for our research. Figure 2 outlines the DSR approach we took and shows the activities we performed in each cycle.

Cycle #1: We began the first cycle with a literature review on data trustees and interviews with urban data experts. The findings led to the design objectives for our prototype (see Section 4.1). We then developed an initial system architecture consisting of the software components required to meet the design objectives. In a second step, we implemented the previously defined system architecture using open source components. The first prototype realized key functionalities required for trusted data exchange and policy management.

Cycle #2: In the second cycle, we improved the initial prototype by extending the value-added services functionality for the application scenario, namely the creation of energy certificates for buildings. For this purpose, we implemented functionalities that allow participants to authenticate themselves in the data space and share their data via their data catalog or retrieve data from the counterparty's data catalog.

Cycle #3: The third cycle was about providing a better user experience. Thus, our developments focused on implementing a suitable dashboard, supporting the management of data assets, the distribution of data to other data space participants, and the creation and handling of energy certificates. We describe our final prototype in more detail in Section 4 below.

To assess whether our prototype was achieving its purpose and was practically relevant, we conducted qualitative evaluations at the end of each DSR cycle (Peffer et al., 2007; Venable et al., 2012). The evaluations were organized as focus group discussions involving the DSR team for cycles #1 and #2, and the DSR team with external parties for cycle #3, following established guidelines (Krueger and Casey, 2015). The focus group discussions helped us get feedback on our developments and set the goals for the upcoming DSR cycle.

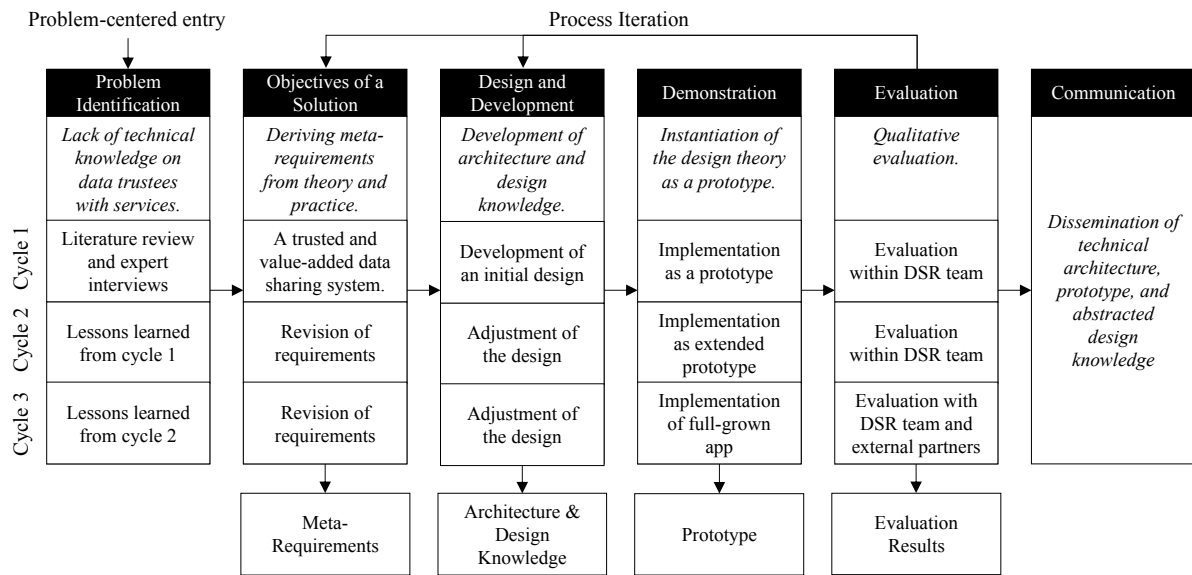


Figure 2: Overview of the DSR Process (adapted from (Peffer et al., 2007)).

4 DESIGNING A PROTOTYPE FOR DATA TRUSTEES

4.1 Design Objectives and Meta-Requirements

Data trustees are trusted intermediaries between data providers and consumers. The DGA outlines the definition and responsibilities of such intermediaries, emphasizing their neutral stance and the prohibition of monetization of the mediated data (European Commission, 2020). In addition, data trustees are tasked with ensuring robust data protection. In the context of building data, stakeholders are also calling for a move from basic data protection to data sovereignty. This requires capabilities to manage data access and usage policies. These allow each data provider to grant or revoke access to data at any time. This results in the following design objective (DO):

DO1a: The Data trustee is an intermediary and empowers building owners by providing them with control mechanisms over data usage and access.

In addition, data is collected from building owners only as needed, not in advance. The data trustee also retains the data only temporarily, ensuring that it is not kept longer than necessary. This approach reduces unnecessary data collection and minimizes security risks, addressing concerns about potential data misuse.

DO1b: Data is stored at the provider's site and accessed only when needed, without creating a centralized data repository.

Furthermore, in the context of building data management, the role of the data trustee goes beyond the mere transfer of data; it also includes acting as a service provider (see (Lauf et al., 2023)) responsible for the creation of the energy certificate. This value-added service (see (Stachon et al., 2023)) is sourced from a decentralized service provider and used by the data trustee to create the energy certificate as a valuable data product that benefits all stakeholders. In addition, it supports broader societal goals such as energy conservation. The service has the potential to drive further innovation and create network effects (Vrabie, 2009). Following objective results:

DO2: The data trustee creates valuable data products that benefit multiple stakeholders, in the form of energy certificates.

In addition to enabling isolated services for a strictly defined set of partners, the data trustee needs to be prepared for data ecosystem contexts. This includes complex service chains and changing stakeholders. This requirement implies:

DO3: The data trustee enables defined interfaces and processes that enable new data consumers, data providers, and service providers in data ecosystems.

In line with the goals of the ecosystem, competition among value-added service providers is welcomed. This is intended to improve the quality and diversity of services available to building owners (Engert et al., 2022). These efforts lead to the final objective:

DO4: The value-added service of the data ecosystem can be enhanced by the inclusion of services pro-

vided by other service providers.

Figure 3 visualizes the data trustee concept resulting from the design objectives. The concept includes the relevant stakeholders of building owner, municipality and energy certificate service as well as possible other data and service providers from the ecosystem. The building data trust acts as a service provider for both intermediary and value-added services, which in our prototype is the energy certificate service, and the ability to adapt other services. It is equipped with interfaces to users and service providers.

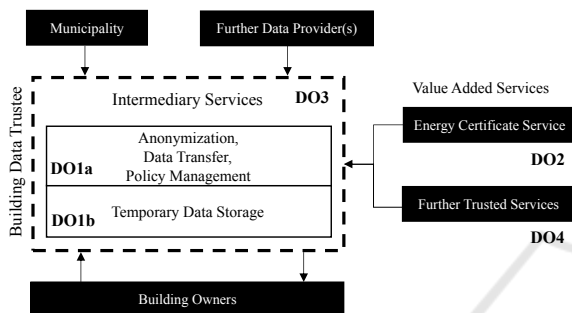


Figure 3: Building Data Trustee Concept Related to Design Objectives.

4.2 Design and Development: System Architecture

The architecture in Figure 4 outlines the structural components that form the foundation of the Building Data Trustee prototype. The architecture describes three distinct areas: First, the energy certificate provider's system, which includes data transfer and a value-added service. Second, the municipality's system, which provides the building owner's consumption data. Third, and most importantly, there is a system for building owners. This system connects the building owners via a user interface with the data trustees and thus with the other parties, the municipality and the energy certificate providers. The data flow is also shown in Figure 4, which outlines the trusted data flow with thick arrows, while the thin arrows visualize the preceding metadata flows. The metadata flows ensure that the data is only transferred when all conditions are met.

The prototype uses data space technologies to enable secure and trusted data transfer. These technologies provide essential features required by data trustees, such as secure data exchange and policy management. A key component of data spaces is the connector that facilitates these processes.

Connectors are extensible gateways for managing data transfers based on metadata and policies. The

prototype architecture has three connectors. They provide the operational channels through which data flows between the building owner, the municipality and the energy certificate provider. Each connector is used as part of the data exchange process, ensuring that data is transferred securely and according to pre-defined agreements.

To further enhance privacy, we have developed a connector extension that anonymizes structured data before it is transferred to the data sink. This ensures that sensitive data remains protected throughout the exchange process, meeting privacy and compliance requirements while maintaining data integrity.

As part of the implementation of the prototype, we are using Eclipse Dataspace Components¹ as a potential framework for the Dataspace Protocol² (DSP), which is governed by the International Data Spaces Association³ (IDSA). This provides an open-source and standards-based approach to building the connectors, ensuring adherence to widely accepted data sovereignty and interoperability standards.

The EDC framework implies interoperability because any system that implements its connector can communicate and exchange data with other connectors that follow the DSP. This ensures seamless data exchange and integration between different systems, increasing the efficiency and effectiveness of data flows within the ecosystem. Scalability in the EDC framework is facilitated by the conceptual separation into a management plane and a data plane. The management plane is used to prepare the necessary metadata with information about the data to be exchanged and the policies, and to negotiate the data transfer. First, the management plane checks all conditions for data sharing. The data plane then executes the data transfer. The data plane is horizontally scalable, meaning that many instances of the data plane can run simultaneously to accommodate increasing data demands. This design allows resources to scale efficiently to meet the needs of growing data volumes and the number of building owners, ensuring that the architecture can adapt and expand without compromising performance or security.

Another component of the architecture is the identity hub, which is managed by each connector. This hub handles the authentication and authorization processes. It verifies the identities of all participants in the ecosystem, ensuring that only authenticated and authorized entities can engage in data transfers. By maintaining a high level of security and trust, the

¹<https://projects.eclipse.org/projects/technology.edc>

²<https://docs.internationaldataspaces.org/dataspace-protocol>

³<https://internationaldataspaces.org>

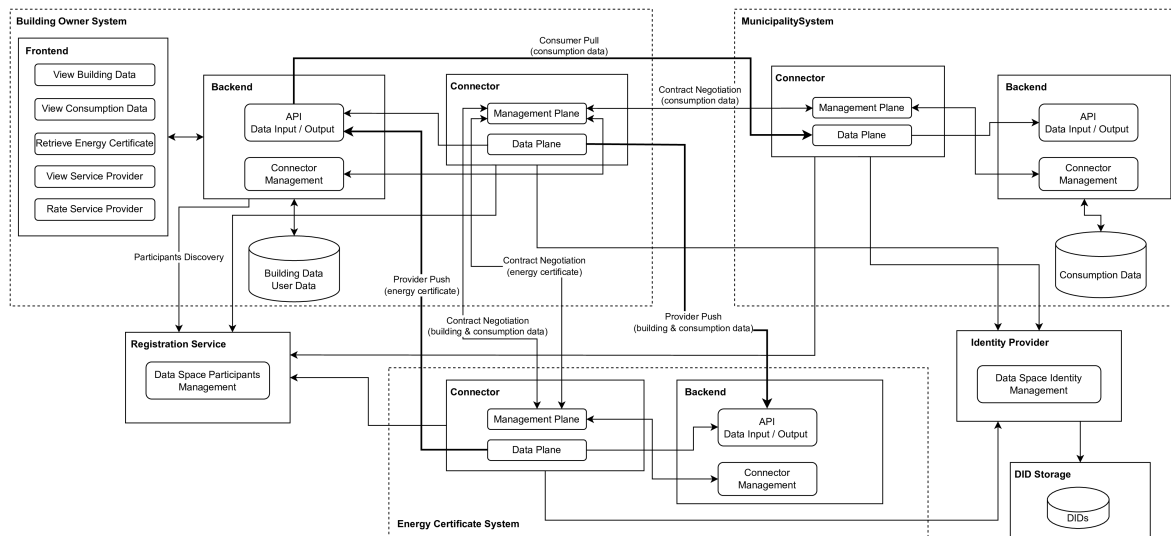


Figure 4: Architecture of Building Data Trustee Prototype.

identity hub plays a critical role in protecting sensitive data environments.

In addition, the registration service acts as a central directory that allows connectors to discover each other. This allows them to expand their data catalogs by discovering and leveraging the data offered by other connectors. This centralized registration ensures that all participants in the data space are properly cataloged and that new participants can be seamlessly integrated into the existing ecosystem. This organization of data flows ensures that the exchange between all participants is seamless.

The identity provider component is responsible for managing participant identities throughout the ecosystem. Using decentralized identifiers (DIDs) stored by participants, the identity provider assigns each connector a unique identity within its identity hub, along with associated claims in the data space. These identities and claims allow connectors to prove both their authenticity and their authorization to exchange data. By verifying these identities and claims, connectors can ensure that only authorized participants are accessing the data intended for them. The identity provider, operated by a central and trusted authority within the data space, thus guarantees the authenticity, integrity, and security of the entire ecosystem.

To facilitate user access, the prototype includes both a frontend and a backend for user management. This backend is essential because the connector itself is not multitenant – meaning that without fine-grained access control, users could theoretically access all data. The backend for user management mitigates this risk by assigning specific data products to individual users, such as building owners, ensuring

that they only have access to the data that is intended for them. In addition, the frontend and backend act as an abstraction layer, simplifying interactions with the data space. Through an intuitive dashboard, users can perform various tasks – such as viewing building data, consumption data, service providers, retrieving energy certificates, or rating service providers – without having to understand the technical complexities of the data space. The backend manages the intricate processes of data management and interaction within the data space, ensuring a seamless user experience.

We selected the EDC framework because it provides a robust open source implementation of emerging data space standards (i.e., DSP), which aligns with our goal of ensuring interoperability and ecosystem readiness. The modular architecture and adherence to the DSP simplify integration with multiple services and stakeholders, reducing time to market and fostering vendor-neutral collaboration. By building on the EDC framework, we leverage established, community-driven components rather than developing proprietary solutions from scratch, ensuring maintainability and compliance with evolving data sharing regulations. This strategic choice supports our goals of data sovereignty, security, and extensibility in a decentralized and dynamic data ecosystem.

4.3 Prototype Implementation

The prototype as a data trustee facilitates the interaction between building owners, the municipality and the certificate provider. It supports the creation and negotiation of data contracts according to the specifications of the data space used, thus enabling the secure and sovereign exchange of consumption and

building data for the generation of energy certificates. In addition to the use of decentralized services, another role of the data trustee could be the certification of these services. One possible implementation within the data space would be for the data trustee to store a hash value, derived from the test data used during the certification process, within the metadata description of the service in the data catalog. This ensures integrity by making it obvious if the service provider is delivering a manipulated service that differs from the one that was verified and cataloged in the metadata.

This practical application demonstrates the translation of theoretical architectural aspects into a functional system and demonstrates the potential of the prototype as a data trustee for real-world applications.

| Building ID | Flat Number | Last Name | Address | Living Space | Rooms | Water Meter | Appliances | Service Provider | Rating | Energy |
|-------------|-------------|-----------|--------------|--------------|-------|-------------|------------|------------------------|--------|--------|
| 11000001 | 101 | Max | Munichstr. 1 | 71 | 10 | 1 | 2 | Local Service Provider | 4.5 | 1.10 |
| 11000002 | 102 | Max | Munichstr. 1 | 71 | 10 | 1 | 2 | Local Service Provider | 4.5 | 1.10 |

Figure 5: Prototype view of a building owner's managed buildings.

Figure 5 shows the building overview of a building owner from our prototype. The user journey is as follows: The building owner registers his building in the system 1). He can then edit 1.1) or delete 1.2) his building data. If his building data is valid, he can retrieve his consumption data from the municipality 2). This process involves interaction with the data space in the background, where the municipality creates an asset in its data catalog that represents the building owner's consumption data. This asset is assigned a access policy that allows only the building owner to view and query this data in the municipality's data catalog.

A data contract is then negotiated between the municipality and the building owner based on the asset and its initial usage policy. In this negotiation, both parties agree on various aspects of the data usage, such as access rights, duration, and scope. Once both parties reach an agreement, the consumption data is transferred.

An existing data contract can also be terminated, which ends the data exchange between the parties. Once the contract is terminated, the building owner loses access to the municipality's consumption data.

After gaining access to their consumption and building data, the building owner can select the cer-

tificate provider 3) as a service provider and use its service to create an energy certificate for their building 4). To do this, the certificate provider's service is again securely exchanged in the data space. Once the service is exchanged, the data trustee can generate the energy certificate and the building owner can display the energy certificate 5).

4.4 Evaluation

For evaluation, the design objectives are mapped to the prototype implementation.

DO1a aimed to give building owners control over their data. This was achieved by enabling secure access to their consumption data from the municipality for various services, including the generation of energy certificates. A user-friendly dashboard was developed to give building owners control over their data sharing preferences and access policies, significantly improving privacy and trust in the services offered.

DO1b is realized by using the data space connector, a system that manages data flows based on metadata on each participant's side.

To address *DO2*, we have implemented an energy certificate service and established a secure and standardized data flow between building owners and service providers.

DO3 is supported by open source components. The implementation of open APIs was crucial for service integration and system extensibility, laying the foundation for future applications and network effects. This achievement underscores our commitment to enabling the creation of valuable data products, promoting interoperability and unlocking the potential for innovation.

To meet *DO4*, the prototype included a detailed overview of service providers, allowing building owners to compare existing ratings and select services based on their needs. By providing service ratings and feedback, transparency is maintained for informed decisions and healthy competition is encouraged. This effort aims to improve service quality and diversity through competitive dynamics, providing benefits to stakeholders within the ecosystem.

4.5 Operating Model

We have considered the following two operating models for our prototype data trustee.

Operating Model 1: Data Trustee as Operator of the Data Space

The data trustee acts as the creator and operator of the data space, which provides the underlying infras-

structure. In this central role, the data trustee takes on the administration of attribute-based access rights by issuing verifiable credentials (VC), which service providers receive after a check in order to be able to use anonymized or aggregated user data. In return, the service providers offer their services, which enables data to be exchanged for services.

There is a clear advantage for service providers: they gain access to a growing selection of privacy-friendly services without compromising their privacy. At the same time, service providers can obtain valuable data to optimize their services. The data trustee acts as a central trust authority and benefits by providing the infrastructure and controlling the transactions between service providers and service providers.

The central role of the data trustee as the data space operator brings with it potential risks of centralization. This could lead to concerns about a single point of failure or concentration of power, which could compromise the reputation of the entire ecosystem. To minimize these risks, the data trustee can implement clear and transparent governance mechanisms based on regular reviews and the involvement of relevant stakeholders. These structures make it possible to make decisions comprehensible and prevent the concentration of power. In addition, independent audits of the infrastructure and security protocols are carried out to identify potential weaknesses at an early stage. This could be carried out by third parties that are trustworthy for both building owners and service providers.

Operating Model 2: Data Trustee as a Participant in the Data Space

The data trustee acts as a trusted participant within an existing data space managed by an external operator. The data trustee provides services primarily focused on the protection, anonymization, and aggregation of building owner data. As a data intermediary, the data trustee facilitates the secure exchange of anonymized data between building owners and service providers in the data space.

Service providers benefit by receiving anonymized and aggregated data that helps them improve and tailor their services without compromising user privacy. In return, service providers pay the data trustee for access to the data, creating a revenue stream. Meanwhile, building owners retain control of their data and gain access to personalized and privacy-preserving services without revealing their identity.

5 CONCLUSIONS

Our prototype demonstrates the importance of balancing security and usability. By leveraging open source components and adhering to established de facto standards, we have ensured secure and sovereign data transfer without having to build an infrastructure from scratch. This foundational software infrastructure has proven essential, demonstrating that enhanced data security can be achieved without sacrificing usability or efficiency.

The implementation of the Self-Sovereign Identity concept and the use of DIDs were key to ensuring data sovereignty and privacy, as well as the decentralized approach. This approach ensured that users remained in control of their personal information and aligned our prototype with privacy standards. Interoperability and adherence to standards were key to harmonizing integration across different service providers, highlighting the need for a consistent method to ensure the smooth operation of services. Our experience has shown that the choice of data held is a critical factor in balancing security and efficiency. For example, while personal data, such as building data, requires the highest security standards, a simpler and more cost-effective solution may be sufficient for less sensitive data. The scalability of the prototype and its interoperability across service providers has been instrumental in achieving this balance, laying the foundation for a system that can adapt to the growing and evolving needs of building owners and service providers alike.

With an emphasis on ecosystem readiness, we have also enabled the easy integration of additional service providers through the modular and extensible software. This strategic move not only expands the range of services within our data trustee, but also fosters a dynamic and scalable ecosystem.

Our prototype was designed with an inherent openness to allow extensions for a variety of purposes and to include a wide range of data providers and consumers. The principle of openness invites continuous development and collaboration, ensuring the adaptability and flexibility of the prototype to meet the real needs of its users.

A key focus in the design of our prototype was to enhance the user experience and ensure easy accessibility. To achieve this, we prioritized the creation of an intuitive dashboard that simplifies data management and interaction processes for users. We wanted to make the benefits of our prototype easily accessible to all users, regardless of their level of technical expertise.

6 LESSONS LEARNED

The lesson from our prototype is that technology alone is not sufficient to ensure compliance with acceptable usage policies. Therefore, our implementation must be complemented by a robust legal framework that mandates legal compliance for post-transfer data use. In addition, our prototype builds trust through a multi-layered approach, first by leveraging open source technologies that provide transparency and community-driven security, and then by implementing a user-centric rating system that evaluates service providers based on their performance and reliability. This approach ensures that trust is not blindly given, but earned and maintained. Another observation was the importance of a decentralized structure to prevent the undue accumulation of power within a single service provider. To this end, the implementation of rotation mechanisms and incentive systems is essential to prevent the formation of super service providers and to promote a competitive environment among multiple service providers.

Feedback and continuous improvement have been critical to the development of our prototype, allowing us to identify and address usability issues early on. This feedback loop has led to iterative improvements that have increased the functionality and usability of the prototype. However, we anticipate technical challenges, particularly in integrating disparate data sources and ensuring data consistency and quality. Overcoming these challenges will require ongoing collaboration and engagement with all ecosystem participants. Finally, user acceptance and trust will depend heavily on the transparency and accountability of processes within the data trustee. It is therefore important to develop processes that build trust. This can be achieved, for example, through the graphical visualization of data lineage and processing, which provides users with a clear view of the processing and flow of their data.

7 CONCLUSION AND OUTLOOK

Our prototype demonstrates novel solutions for data sovereignty and privacy in the digital economy in a communal environment. It contributes to the scientific community by demonstrating the practicality and effectiveness of data trustees. The implementation of the prototype provides a valuable contribution to the understanding of how data trustees can be designed to balance the interests of various stakeholders in a trustworthy manner. The progress made in developing the prototype illustrates the potential of standardized,

open architectures and underscores the importance of open sources and communities. By making our extensions to the EDC framework available as an open source project⁴, we have actively contributed to the data exchange community and promoted a practical approach to the evolution of data spaces.

However, implementing data trustees in practice requires navigating a complex regulatory environment. The evolving regulatory landscape, including the DGA and GDPR, is challenging the widespread adoption of data trustees. Complying with various national, regional, and domain-specific regulations and aligning stakeholder interests can create barriers to scale. Moreover, user adoption remains uncertain; building owners may resist the data trustee due to privacy concerns, mistrust, or perceived complexity of the technology. Ongoing stakeholder engagement is required to ensure compliance, trust, and adaptability as regulations and expectations evolve.

Future research directions could focus on refining the technical enforceability of usage policies, further developing decentralized structures to avoid monopoly positions, and deepening mechanisms to enhance user trust. It is also important to study the impact of these technologies on different industries and to explore the applicability of our findings to other areas of the digital economy.

As we develop design principles in future work, we will ensure that no design principles (DPs) are missing that would cause significant phenomena in the environment (deficit) and that no additional DPs trigger irrelevant phenomena (excess). Furthermore, we will ensure that a DP does not handle multiple phenomena (redundancy) and that no more than one DP addresses each phenomenon (overload). This type of evaluation has been applied to design principles in previous research (Janiesch et al., 2020), and we have adapted this method for our evaluation. We consider DPs to be a concrete form of design knowledge, and as such we plan to further develop and evaluate these principles in future work. With these additional insights, we intend to formulate a design theory for data trustees in future studies.

Promoting data sovereignty and privacy remains an ongoing challenge in the digital world. Our data trustee prototype is a step towards addressing these challenges, offering novel solutions that are secure, efficient, user-friendly and transparent. Beyond the communal businesses, our data trustee empowers individual building owners by protecting their data sovereignty and ensuring their privacy.

⁴<https://github.com/MichaelSteinert/edc-anonymize-http-data-plane>; <https://github.com/MichaelSteinert/edc-rating-participant>

ACKNOWLEDGMENTS

The authors acknowledge the financial support by the Federal Ministry of Education and Research (BMBF) of Germany in the project KomDatIS (grant number: 16DTM106A-C).

REFERENCES

- Blankertz, A. and Specht-Riemenschneider, L. (2021). Neue modelle ermöglichen – regulierung für daten-treuhänder.
- Bogers, M., Sims, J., and West, J. (2019). What is an ecosystem? incorporating 25 years of ecosystem research. *SSRN Electronic Journal*.
- Curry, E. (2016). *The big data value chain: definitions, concepts, and theoretical approaches*. Springer International Publishing.
- Curry, E. (2020). *Real-time Linked Dataspace: Enabling Data Ecosystems for Intelligent Systems*. Springer Nature.
- EnergyREV, editor (2020). *Privacy and data sharing in smart local energy systems: Insights and recommendations*. University of Strathclyde Publishing, Glasgow, UK.
- Engert, M., Evers, J., Hein, A., and Krcmar, H. (2022). The engagement of complementors and the role of platform boundary resources in e-commerce platform ecosystems. *Information Systems Frontiers*, 24(6):2007–2025.
- European Commission (2020). Regulation of the european parliament and of the council on european data governance: (data governance act).
- European Commission (2022). The value of energy data and its role in the market.
- European Union (2010). Directive 2010/31/eu of the european parliament and of the council.
- Federal Ministry of Education and Research Germany (2022). Datentreuhandmodelle: Pilotvorhaben des bmbf.
- Hardinges, J., Wells, P., Blandford, A., Tennison, J., and Scott, A. (2019). Data trusts: lessons from three pilots.
- Heuninckx, S., Meitern, M., te Boveldt, G., and Coosemans, T. (2023). Practical problems before privacy concerns: How european energy community initiatives struggle with data collection. *Energy Research & Social Science*, 98:103040.
- Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1):75–105.
- Janiesch, C., Rosenkranz, C., and Scholten, U. (2020). An information systems design theory for service network effects. *Journal of the Association for Information Systems (JAIS)*, 21:1402–1460.
- Jenks, M. and Jones, C., editors (2010). *Dimensions of the Sustainable City*, volume 2. Springer Netherlands, Dordrecht.
- Krueger, R. A. and Casey, M. A. (2015). *Focus groups: A practical guide for applied research*. SAGE, Los Angeles and London and New Delhi and Singapore and Washington DC, 5th edition edition.
- Lauf, F., Scheider, S., Friese, J., Kilz, S., Radic, M., and Burmann, A. (2023). Exploring design characteristics of data trustees in healthcare-taxonomy and archetypes. In *ECIS 2023 Research Papers*, volume 323. AIS Electronic Library (AISeL).
- Li, Y., Kubicki, S., Guerriero, A., and Rezgui, Y. (2019). Review of building energy performance certification schemes towards future improvement. *Renewable and Sustainable Energy Reviews*, 113:109244.
- Moore, J. F. (1993). Predators and prey: a new ecology of competition. *Harvard Business Review*, 71(3):75–86.
- Moran, D., Kanemoto, K., Jiborn, M., Wood, R., Többen, J., and Seto, K. C. (2018). Carbon footprints of 13 000 cities. *Environmental Research Letters*, 13(6):064041.
- Oliveira, M. I. S. and Lóscio, B. F. (2018). What is a data ecosystem? In Janssen, M., Chun, S. A., and Weerakkody, V., editors, *Proceedings of the 19th Annual International Conference on Digital Government Research Governance in the Data Age - dgo '18*, pages 1–9, Delft: Netherlands. ACM Press.
- Otto, B. and Burmann, A. (2021). Europäische dateninfrastrukturen. *Informatik Spektrum*, pages 1–9.
- Peppers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3):45–77.
- Porru, S., Misso, F. E., Pani, F. E., and Repetto, C. (2020). Smart mobility and public transport: Opportunities and challenges in rural and urban areas. *Journal of Traffic and Transportation Engineering (English Edition)*, 7(1):88–97.
- Stachon, M., Möller, F., Guggenberger, T. M., Tomczyk, M., and Henning, J.-L. (2023). Understanding data trusts. In *ECIS 2023 Research-in-Progress Papers*, volume 36. AIS Electronic Library (AISeL).
- United Nations (2015). Transforming our world: The 2030 agenda for sustainable development.
- Venable, J., Pries-Heje, J., and Baskerville, R. (2012). A comprehensive framework for evaluation in design science research. In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Peppers, K., Rothenberger, M., and Kuechler, B., editors, *Design Science Research in Information Systems. Advances in Theory and Practice*, volume 7286 of *Lecture Notes in Computer Science*, pages 423–438. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Vrabie, C. (2009). Just do it – spreading use of digital services. In *EPGA Conference 2009*. IEEE Computer Society.