Vasileios Karagiannis, Georg Simhandl, Dražen Ignjatović, Bernhard Bürger, Mario Drobics, "The Role of Data Spaces in the Internet of Things", *Springer e+i Elektrotechnik und Informationstechnik*, 2025, DOI: https://doi.org/10.1007/s00502-025-01320-6

The Role of Data Spaces in the Internet of Things

Vasileios Karagiannis, Georg Simhandl, Dražen Ignjatović, Bernhard Bürger, Mario Drobics

Center for Digital Safety & Security, Austrian Institute of Technology, Giefinggasse 4, 1210, Vienna, Austria.

*Corresponding author(s). E-mail(s): vasileios.karagiannis@ait.ac.at;
Contributing authors: georg.simhandl@ait.ac.at; drazen.ignjatovic@ait.ac.at;
bernhard.buerger@ait.ac.at; mario.drobics@ait.ac.at;

Abstract

The Internet of Things (IoT), empowered by edge and cloud computing, has driven innovation across numerous sectors by enabling real-time data processing and decision-making. However, sectors such as crisis management, energy, and manufacturing can be hindered by the stringent requirements of handling highly sensitive data. To foster innovation in such sectors, modern regulations propose advanced digital services and functionalities to manage sensitive data in a secure, trusted, and transparent manner. For example, the General Data Protection Regulation (GDPR) governs the handling of personal data, the Data Act and the Data Governance Act facilitate data sharing with third parties and intermediation services, and the electronic IDentification, Authentication and trust Services (eIDAS) regulate electronic authentication. Nevertheless, navigating these regulations and designing compliant IoT systems can pose challenges due to the diversity of the stakeholders, the sensitivity of the data, and the criticality of the applications. To tackle such challenges, we propose the integration of data spaces into IoT systems. This approach relies on data policies to share and process data in alignment with regulations, enabling innovation in sectors dealing with highly sensitive information. Moreover, we identify the associated challenges, and we propose future research directions to further advance the IoT.

Keywords: Edge computing, Computing Continuum, Internet of Things, Data Spaces, Crisis Management, Energy, Manufacturing

1 Introduction

So far, the IoT has revolutionized modern societies by enabling real-time data processing and analytics which contribute to informed and often automated decision-making [1]. Due to such advancements, various sectors have been significantly improved with innovative applications in transportation, smart homes, agriculture, and fitness, among others. The deployment of such applications typically spans multiple computing layers, including IoT, edge, fog, and cloud, which aligns with the paradigm of the computing continuum [2]. The computing continuum integrates diverse computing resources to create a fluid and scalable network, enhancing the performance and capabilities of IoT systems. Consequently, novel IoT applications typically operate over the continuum in order to provide high performance, resilience, and scalability.

However, some sectors have not yet experienced considerable IoT innovations due to rigorous regulatory compliance requirements and strict data protection laws that might impede the digital transformation of the IoT [3]. In crisis management, for example, there are privacy concerns surrounding the handling of personal information during emergencies, which may require compliance with GDPR. Similarly, the energy sector faces stringent requirements, especially due to

its designation as critical infrastructure, which necessitates high standards of data security and integrity under the NIS2 Directive. Furthermore, in the manufacturing sector, extensive data collection about manufacturing processes might be required under the Registration, Evaluation, Authorization, and Restriction of Chemicals (REACH) regulation, and the Emissions Trading System (ETS) demanding accurate monitoring and reporting of greenhouse gas emissions for managing and reducing the carbon footprint. In addition, data from these sectors may need to be shared in a secure and trusted manner as stipulated by the Data Governance Act. Thus, legal and regulatory requirements can complicate the handling of IoT data [4], which can limit the innovative transformation of the IoT in such sectors.

To avoid this limitation, this paper introduces the concept of the data space as a data management system that operates in alignment with legal regulations. This alignment is achieved by using data policies, formulated using the Open Digital Rights Language (ODRL), to represent applicable regulations when sharing data [2]. Data policies complement the sharing of data, ensuring that all crucial information about data-handling regulations is conveyed during the data exchange. The proposed approach also relies on local edge computing resources to enable secure end-to-end encrypted data sharing directly between system participants. This aims to prevent the use of third-party storage (for example, via commercial cloud services) that might violate the strict privacy requirements of the data, e.g., due to customer data analytics [5]. Furthermore, using edge resources to build data spaces has been shown to result in low data-sharing latency [6]. Therefore, using a combination of edge computing and data spaces, we aim at overcoming the limitations of sectors dealing with highly sensitive data. Specifically, we address the three identified sectors with strict regulatory requirements, namely, crisis management, energy, and manufacturing. For these sectors, we discuss tailored data sharing systems that are able to adhere to regulations and enable innovative IoT applications.

Overall, the following prime contributions are within the scope of this paper: We identify challenges in extending the IoT across sectors with strict regulatory requirements. We propose the integration of data spaces and edge computing to address these challenges, and we present tailored data sharing systems for three specific sectors. Finally, we discuss open challenges and future research directions with the potential to aid the implementation of compliant data sharing in the IoT.

The remainder of this work is structured as follows: Section 2 introduces the concept of data spaces and discusses related work from the literature. Afterward, Sections 3, 4, and 5 present data spaces for the sectors of crisis management, energy, and manufacturing, respectively. Subsequently, Section 6 proposes Gaia-X for facilitating trust and interoperability between data spaces of different sectors, and Section 7 focuses on the associated challenges. Finally, in Section 8, we conclude this paper and propose future work on this topic.

2 Background and Related Work

Data spaces are emerging as the standard approach to enable secure and compliant data sharing. A data space is a data management system that governs data access and usage through policies, creating a trusted environment where stakeholders within a sector can exchange information without losing data ownership privileges. A data policy is a set of rules, formulated using ODRL, that define how the data can be accessed, used, and shared, representing regulatory and organizational requirements. When sharing data through a data space, the data consumer must first agree to the policy in order to access the data. This ensures that the data provider, who defines the data policy, can share information without losing control over the data. All data space participants are verified organizations that use the data space to share valuable information while maintaining access and usage control. This fosters trust among participants and facilitates information sharing, even between stakeholders who would be reluctant to share their data otherwise. To enable such functionalities, a data space includes various software components including catalogs/registries for metadata and semantic models, identity providers, and dataspace connectors [6].

Various works in the literature present data spaces, and discuss the associated benefits. In the context of the computing continuum in general, data spaces improve sovereignty and compliance [2]. The same applies considering specific use cases and sectors. For example, the Blue Dataverse presents a technical approach to build a data space for environmental monitoring, enabling data analytics related to climate change and sustainability [7]. BatWoMan introduces a data space for battery passports, and provides a technical architecture for decentralized data exchange [8]. In addition, the Gaia-X Hub Austria has produced a technical overview of data spaces, explaining existing software tools and solutions [9]. When using edge computing to build data spaces, early work has explored

various system architectures [10], as well as latency and bandwidth benefits [6]. To further advance the state of the art of data spaces and the IoT, this paper focuses on specific sectors with strict regulatory requirements, which are mostly unexplored in the literature.

A concrete example of data sharing using a data space, e.g., in the energy sector, may include stakeholders such as customers, utility companies, and energy service providers. A utility company collects metering data about the energy consumption of the customer from smart meters at the edge of the network, which is primarily used for billing. Since this data is private, it can be subject to regulations, so it is stored in the private edge computing infrastructure of the utility provider, and is not shared with third parties unless explicitly requested by the customer. Utility providers typically offer a web interface for customers to access their data. However, this web interface may provide limited functionality, e.g., basic visualizations, because utility providers do not commonly specialize in data science and analytics. External energy service providers, on the other hand, specialize in energy analytics and optimizations, but do not have access to customer data by default. Nevertheless, when both the utility provider and the energy service provider join a data space, sharing customer data becomes streamlined. In this case, the utility provider can offer external analytics services on the web interface, that are executed by external service providers. To use an external service, the customer fills out a consent form on the web interface, allowing the service providers to process the customer data. Based on this form, the utility provider creates the data policy, and shares the customer data (and the policy) with the service provider using dataspace connectors. The service provider processes the data locally on edge resources, and returns the results to the utility provider that shows the results to the customer through the web interface. This approach enables third-party service providers to process customer data in compliance with regulations, while having the policy that can be used as proof of consent. The data is transferred end-to-end encrypted to ensure confidentiality, and the policy can be digitally signed by all parties for accountability. Thus, exchanging data in this manner using edge resources and data spaces can enable compliant data sharing for the energy sector, and similarly, for other sectors as well.

3 Data Space for Crisis Management

A data space for crisis management can provide access to critical information, such as resource availability, emergency response capabilities, and affected populations. This information can be crucial, especially for timely decision-making [11] across all the phases of a disaster: prevention, preparedness, response, and recovery. By accessing and processing this information, stakeholders can utilize various applications to coordinate disaster risk reduction efforts, such as: resource coordination which allocates personnel and equipment effectively, situation assessment which processes data to understand the scale and impact of an incident, and resilience planning which involves analyzing data to enhance preparedness and infrastructure resilience. Therefore, sharing crisis management data among the stakeholders can provide various benefits.

Utilizing a data space to share information among stakeholders in crisis management can address key challenges that are otherwise difficult to overcome. Crisis management often involves sharing personal data of affected populations (e.g., location, contact, and health records), sensitive information about critical infrastructure (e.g., power or water supply), and situational reports (e.g., about the status of ongoing operations or evolving weather conditions). Sharing personal data may need to comply with the GDPR, which often requires explicit user consent for processing health-related data. Likewise, sharing data about critical infrastructures may need to align with the NIS2 Directive. Furthermore, since disasters can transcend national borders, stakeholders might need to account for diverse regulations from different countries and always process the data in compliance with the local regulations. This means that different stakeholders may process the data differently based on their country. Interestingly, data treaties between countries may ensure that the regulations of the data's country of origin are respected, although the concept of a data treaty is not yet widely accepted. These requirements may need to be met to allow for collaborative data processing, which can be vital for effective crisis response and risk reduction. A crisis management data space can address such requirements by providing data along with associated policies, thereby communicating applicable regulations to stakeholders. Additionally, data spaces can resolve existing issues like incompatible data formats, data silos, and access control barriers, which often delay the decision-making process. Thus, a data space for crisis management, as shown in Figure 1, can provide a system whereby various

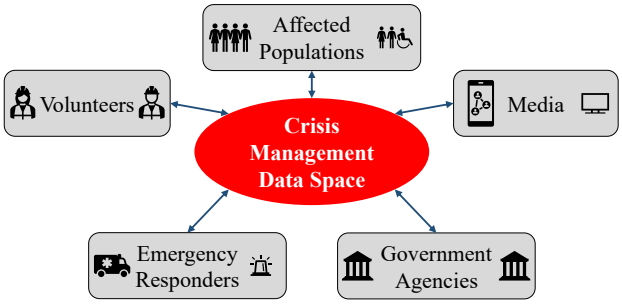

Fig. 1 Stakeholders of the crisis management sector sharing data through the data space.

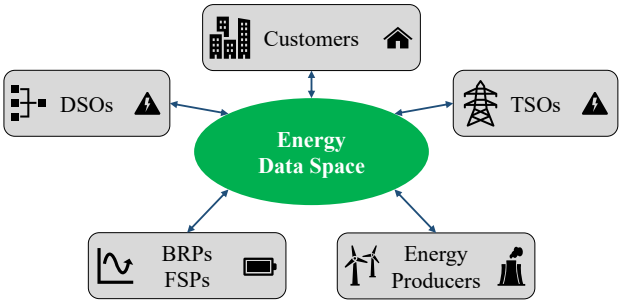

Fig. 2 Stakeholders of the energy sector sharing data through the data space.

stakeholders such as affected populations, emergency responders, decision-makers, volunteers, media outlets, government agencies, and infrastructure operators share data in a secure, transparent, and compliant manner. Furthermore, integrating edge computing can facilitate direct data exchange between participants, enhancing data privacy and availability, and reducing delays [12].

4 Data Space for Energy

A data space for energy can provide secure and compliant access to energy data [13]. Energy data refers to information related to energy production, distribution, and consumption, including power grid load, energy consumption patterns, and renewable energy integration. This data is crucial for ensuring the efficient operation, optimization, and sustainability of energy systems and power grids, which are foundational for modern societies. By accessing and processing energy data, stakeholders can utilize various applications regarding, for example, smart grids which aim at balancing supply and demand, predictive repairs which restore faults before they cause failures, and demand-side management which allows customers to optimize energy use and reduce costs.

Utilizing a data space to share data across stakeholders in the energy sector can address key requirements that cannot be easily met by traditional data sharing approaches. For example, customer energy consumption information may qualify as personal data under the GDPR. Thus, utility companies collecting this data, e.g., for billing purposes, might have to adhere to privacy principles, including data minimization and purpose limitation. Additionally, under the Data Act, customer data may need to be made accessible to third parties like energy service providers that enable digital services offering energy visualization and optimization features. Furthermore, the energy sector is classified as critical infrastructure under the NIS2 Directive, requiring high security and integrity. When sharing energy data using a data space, such requirements can be formulated (into policies) and shared along with the data so that stakeholders can access and process this information accordingly. A data space can also implement role-based access control, ensuring that verified organizations access the data based on their roles. For instance, utility companies may have access to sensitive infrastructure data that is not accessible to other organizations or customers. Moreover, integrating edge computing and keeping the data distributed on local organizational storage ensures privacy, availability, and low-latency processing [14]. This way, the stakeholders of the energy sector, including customers, Transmission System Operators (TSOs), Distribution System Operators (DSOs), Flexibility Service Providers (FSPs), Balance Responsible Parties (BRP), energy producers, and energy service providers, also shown in Figure 2, can share data in a secure, transparent, and compliant manner. This enables collaborative data processing and advanced optimization across energy production, distribution, and consumption, which might not be possible by optimizing each layer separately.

5 Data Space for Manufacturing

A data space for manufacturing can provide access to data about industrial processes, such as: emissions, resource utilization, product specifications, and reporting, among others. This data is essential for optimizing production efficiency, ensuring regulatory compliance, and supporting sustainability efforts, which are increasingly critical in modern societies. Accessing and processing such data can enable various applications for stakeholders like: predictive maintenance which analyzes machinery data to prevent equipment failures, supply chain optimization that improves delivery timelines and reduces costs, and provenance verification which ensures that materials

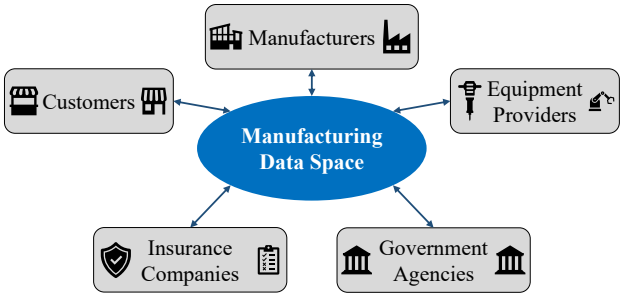

Fig. 3 Stakeholders of the manufacturing sector sharing data through the data space.

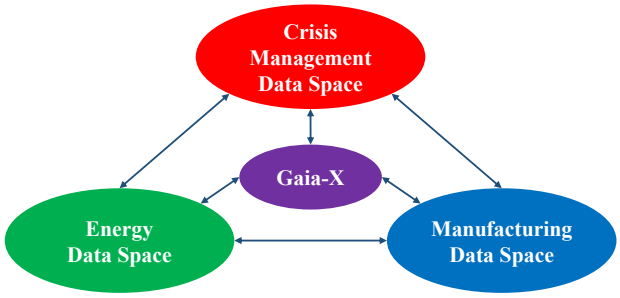

Fig. 4 Different sectors sharing data, supported by Gaia-X.

and processes meet regulatory and sustainability standards. Also, the integration of this data can facilitate better decision-making and foster innovation across manufacturing processes, ensuring both economic and environmental sustainability.

Utilizing a data space to share manufacturing data among stakeholders can satisfy key sector requirements. For example, under REACH, manufacturers may need to share information about chemical processes, the ETS mandates reporting of greenhouse gas emissions, and the Ecodesign for Sustainable Product Regulation (ESPR) introduces a Digital Product Passport (DPP) which requires manufacturers to provide verifiable information on sustainability, circularity, and compliance. However, sharing such data can present challenges due to trade secrets, privacy concerns, and the need to align with regulations like the Data Act. A manufacturing data space can aid in storing and sharing all related data while addressing these challenges. This can be achieved by enabling the sharing of data with associated policies (that detail regulatory and organizational requirements), and by implementing appropriate access control. Furthermore, integrating with edge computing allows manufacturing data to remain distributed across organizational storage systems [12, 15], preserving privacy and sovereignty. Consequently, a data space becomes an enabler of the DPP concept by facilitating access to all product-related data that may be distributed across different organizations in the supply chain. Overall, a manufacturing data space, as shown in Figure 3, enables stakeholders such as manufacturers, customers, suppliers, equipment providers, companies insuring the equipment, and government agencies to share information and execute collaborative processing of data across the supply chain. This approach can optimize the supply chain, improve operational efficiency, ensure regulatory compliance, and enhance the sustainability of the manufacturing sector.

6 Data Spaces and Gaia-X

Gaia-X aims to provide innovative compliance services that facilitate interoperability between data spaces. In line with regulations such as the GDPR, eIDAS, the Data Act, and the Data Governance Act, Gaia-X offers services that foster interoperability, compliance, and trust across different data space systems. By providing compliance-as-a-service, Gaia-X abstracts the problem of implementing dedicated compliance mechanisms (e.g., for policy and identity validation) within a data space, which can reduce the overhead of building data spaces significantly. While the benefits of sector-specific data spaces are examined (in Sections 3, 4, and 5), collaboration between different data spaces can also have advantages. For example, if crisis management stakeholders have streamlined access to energy infrastructure data, hidden electrical hazards can be identified and mitigated proactively during rescue operations. Similarly, manufacturing stakeholders can schedule energy-intensive processes based on energy supply data to avoid abrupt peak demand, and reduce the grid load. Furthermore, with access to crisis management data, the manufacturing sector can dynamically adapt production to support urgent needs, such as producing essential supplies like cloth face masks during pandemics. This aligns with the concept of a Reconfigurable Manufacturing System (RMS) or a Flexible Manufacturing System (FMS), enabled by the collaboration of data spaces. Thus, by supporting cross-sectoral collaboration, as shown in Figure 4, Gaia-X enhances resilience, efficiency, and sustainability in modern societies. Notably, significant progress towards realizing data spaces and data space interoperability is supported by the Gaia-X Hub Austria and the Austrian Institute of Technology [8, 9, 16–18].

7 Open Challenges and Future Directions

While data spaces hold great potential, as previously discussed, several challenges still remain open. Achieving interoperability between diverse data systems and formats can be complex, requiring flexible protocols and semantic data models. Ensuring low-latency data processing might also be challenging, especially when data is distributed across multiple organizations. Moreover, while edge computing provides significant benefits to data spaces, resources at the edge of the network can be limited, and may not replace the large capacities of the centralized cloud approach. Thus, scalability at the edge can be difficult when sharing data end-to-end. Also, the access to such a diverse data management system might allow participants to partially reconstruct confidential datasets from various accessible data pieces. Accordingly, future research can focus on developing appropriate techniques, protocols, and data models to handle sensitive data and achieve interoperability across diverse data systems. Innovations in edge computing technologies are also needed to implement low-latency data processing in distributed edge systems, and cloud storage systems may need to adapt to offer compliant data-sharing as well. Additionally, scalable architectures and optimization techniques need to be explored to enable data spaces to handle large volumes of data efficiently. Finally, appropriate privacy-preserving mechanisms applicable to data spaces have to be investigated to guarantee privacy and increase trust among the participants. Notably, these research directions need to pursue lightweight solutions which do not introduce significant overhead, in order to ensure efficient system operation.

8 Conclusion

This work presents compliant data sharing using data spaces across the IoT, including in sectors such as crisis management, energy, and manufacturing. By addressing key sector requirements and regulations, we discuss how data spaces enhance operational efficiency, resilience, and sustainability. These benefits can have broad societal impacts, including improved disaster response, optimized energy use, and more efficient manufacturing supply chains, while also fostering economic growth through cost reductions and enhanced regulatory compliance. Future efforts can focus on addressing the identified technical challenges, such as scalability and interoperability, to further advance data spaces and the IoT.

References

[1] Karagiannis, V., Chatzimisios, P., Vazquez-Gallego, F., Alonso-Zarate, J.: A survey on application layer protocols for the Internet of Things. Transaction on IoT and Cloud Computing **3**(1), 11–17 (2015)

[2] Karagiannis, V.: Data sovereignty and compliance in the computing continuum. In: International Conference on Future Internet of Things and Cloud (FiCloud), pp. 123–130 (2024). IEEE

[3] El-Gazzar, R., Stendal, K.: Examining how GDPR challenges emerging technologies. Journal of Information Policy **10**, 237–275 (2020)

[4] Gkotsopoulou, O., Quinn, P.: Data protection and privacy issues of the Internet of Things. In: Internet of Things, Threats, Landscape, and Countermeasures, pp. 1–46 (2021). CRC Press

[5] Liu, Y., Soroka, A., Han, L., Jian, J., Tang, M.: Cloud-based big data analytics for customer insight-driven design innovation in SMEs. International Journal of Information Management **51**, 102034 (2020)

[6] Karagiannis, V., Al-Akrawi, A., Hödl, O.: Data sovereignty at the edge of the network. In: International Conference on Fog and Edge Computing (ICFEC), pp. 33–39 (2023). IEEE

[7] Karagiannis, V., Al-Naday, M., De Block, T.: The Blue Dataverse: A system for marine data sovereignty. In: IEEE World Forum on Internet of Things (WF-IoT), pp. 1–6 (2023). IEEE

[8] Siska, V., Al-Akrawi, A., Zackrisson, M.: Building a sustainable battery supply chain with digital battery passports. In: Interdisciplinary Information Management Talks (IDIMT), pp. 347–354 (2023)

[9] Siska, V., Drobics, M., Karagiannis, V.: Building a Dataspace: Technical Overview. Gaia-X Hub Austria (2023)

[10] Kung, A., et al.: High Level Architecture (HLA) Report. Alliance for IoT and Edge Computing Innovation, 1–88 (2024)

[11] Ignjatović, D., Karagiannis, V., Chettakattu, A., Havlik, D., Neubauer, G.: Crisis management in the era of the IoT, edge computing, and LLMs. In:

International Conference on Internet of Things: Systems, Management and Security (IOTSMS), pp. 224–231 (2024). IEEE

[12] Karagiannis, V., Schulte, S.: edgeRouting: Using compute nodes in proximity to route IoT data. IEEE Access **9**, 105841–105858 (2021)

[13] Karagiannis, V., Nagy, B., Jodkowski, A., Kranner, M., Ignjatović, D.: A review of emerging trends in energy data management systems. In: International Conference on Internet of Things: Systems, Management and Security (IOTSMS), pp. 74–81 (2024). IEEE

[14] Karagiannis, V.: Compute node communication in the fog: Survey and research challenges. In: Workshop on Fog Computing and the IoT (IoT-Fog), pp. 36–40 (2019). ACM

[15] Barzegaran, M., Karagiannis, V., Avasalcai, C., Pop, P., Schulte, S., Dustdar, S.: Towards extensibility-aware scheduling of industrial applications on fog nodes. In: International Conference on Edge Computing (EDGE), pp. 67–75 (2020). IEEE

[16] Donsa, K., Kreiner, K., Hayn, D., Rzepka, A., Ovejero, S., Topolnik, M., Ziegl, A., Pfeifer, B., Neururer, S., Kaltenbrunner, S., et al.: Smart FOX–enabling citizen-based donation of EHR-standardised data for clinical research in Austria. Digital Health and Informatics Innovations for Sustainable Health Care Systems, 83–87 (2024)

[17] Gordea, S., Andresel, M., Drauschke, F., Kahle, P.: Transcribathon. eu: AI supporting collaborative transcription and enrichment of historical documents. In: International Conference on Advanced Visual Interfaces (AVI), pp. 1–3 (2024). ACM

[18] Andresel, M., Siska, V., David, R., Schlarb, S., Weißenfeld, A.: Adapting ontology-based data access for data spaces. In: International Workshop on Semantics in Dataspaces (2024). CEUR-WS