

Gieß, Anna; Hutterer, Andreas

Article — Published Version

The Future of Data Management: A Delimitation of Data Platforms, Data Spaces, Data Meshes, and Data Fabrics

Information Systems and e-Business Management

Provided in Cooperation with:

Springer Nature

Suggested Citation: Gieß, Anna; Hutterer, Andreas (2025) : The Future of Data Management: A Delimitation of Data Platforms, Data Spaces, Data Meshes, and Data Fabrics, Information Systems and e-Business Management, ISSN 1617-9854, Springer, Berlin, Heidelberg, Vol. 23, Iss. 4, pp. 971-997,
<https://doi.org/10.1007/s10257-025-00707-4>

This Version is available at:

<https://hdl.handle.net/10419/335114>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



The Future of Data Management: A Delimitation of Data Platforms, Data Spaces, DataMeshes, and Data Fabrics

Anna Gieß^{1,2} · Andreas Hutterer³

Received: 9 September 2024 / Revised: 16 January 2025 / Accepted: 4 April 2025 /
Published online: 16 July 2025
© The Author(s) 2025

Abstract

As traditional data management approaches increasingly become barriers to innovation, emerging modern approaches such as data spaces, fabrics and meshes provide much-needed flexibility. To demonstrate the potential of these modern approaches, we conducted a structured literature review and comparative analysis of 189 contributions to explore the differences between data platforms, data spaces, data fabrics and data meshes, highlighting their unique advantages in terms of governance mode and application area. Throughout our delimitation, the study reveals that data platforms and data fabrics offer centralized structures, ensuring control, accessibility, and streamlined management while decentralized data meshes and data spaces prioritize flexibility. We provide a data management matrix that provides valuable insights for organizations seeking to optimize their data management strategies in the sharing economy.

Keywords Data management · Data platform · Data space · Data mesh · Data fabric

1 Introduction

In the digital era, the exponential growth in the volume, velocity, and variety of data poses profound challenges to traditional data management approaches (Achanta & Boina 2023). Conventional approaches, such as data lakes and data warehouses, often struggle to meet the dynamic demands of contemporary, data-driven

✉ Anna Gieß
anna.giess@isst.fraunhofer.de

Andreas Hutterer
andreas.hutterer@jku.at

¹ Fraunhofer ISST, Speicherstraße 6, 44147 Dortmund, Germany

² TU Braunschweig, Abt-Jerusalem-Straße 7, 38106 Braunschweig, Germany

³ Johannes Kepler University Linz, Altenberger Str. 69, 4040 Linz, Austria

decision-making processes (Nambiar & Mundra 2022). Their inherent limitations, including centralized architectures and schema-on-write approaches, hinder their ability to facilitate real-time data analysis effectively (Giebler et al. 2019). To address these shortcomings, emerging data management approaches structured around collaborative organizations or partnerships have gained prominence (Machado et al. 2021). These approaches promote stakeholder interaction and mark a transition away from rigid data sharing approaches towards architectures tailored to enhance specific business requirements, such as data transparency (Geisler et al. 2022). This evolution underscores the necessity of adopting more flexible, decentralized data management approaches, which align with the growing complexity of the data landscape and the imperative for adaptability in digital infrastructures (Pörtner et al. 2023). Further, data has emerged as a strategic asset, necessitating a reevaluation of its value and utility (Otto et al. 2015). This redefinition emphasizes the development of data products despite persistent challenges in cross-organizational data sharing (Bartelheimer et al. 2022; Beverungen et al. 2022; Gelhaar et al. 2023). Furthermore, the increasing generation of vast quantities of unstructured data highlights the critical need for FAIR (Findable, Accessible, Interoperable, and Reusable) data infrastructures (Scheffler et al. 2022). The FAIR principles provide a foundational approach for advancing both the discoverability and the usability of data, thus ensuring that data management aligns with organizational goals and broader societal objectives (Wilkinson et al. 2016).

Conventional approaches to data management are being recognized more frequently as hindrances to contemporary requirements (Bode et al. 2023). Modern approaches, such as data spaces, data fabrics, and data meshes, provide the necessary flexibility for data across diverse organizational contexts (Franklin et al. 2005; Kuftinova, Maksimychev, et al. 2022). These approaches not only deliver enhanced agility but also align with evolving business practices, addressing the demands of today's dynamic market environments (Kuftinova, Maksimychev, et al. 2022). Additionally, the growing emphasis on data sovereignty and privacy—reinforced by regulatory frameworks such as GDPR—introduces added complexity to data management practices (Truong et al. 2020). Organizations are increasingly requiring nuanced approaches that offer precise control over data while navigating complex legal and operational specifications (Machado et al. 2022a; Truong et al. 2020). This need for enhanced sovereignty is particularly evident in cross-border data operations, where compliance with varying regulatory requirements adds layers of complexity to organizational procedures (Opriel et al. 2024).

This study seeks to clarify the evolving concepts and interrelationships within modern data management approaches, aiming to support organizations in selecting approaches that align with their specific objectives and operational requirements (Sherman 2015). A detailed examination of the current literature uncovers a notable gap in comparative research (e.g., authors mistakenly use various terms synonymously) and practical guidelines for implementing these approaches in different organizational contexts, revealing a significant lack of comprehensive analysis on their real-world applications and implications for data management in different environments (Bergman & Wang 2013). Consequently, this paper not only addresses the limitations of conventional data management approaches but also highlights

the benefits of embracing flexible data approaches (Azad et al. 2020). The findings emphasize the urgency for organizations to adopt contemporary data management approaches to fully capitalize on their data assets and maintain competitiveness in an increasingly data-driven economy. Given the scarcity of empirical studies on the adoption and effectiveness of these approaches, this research aims to support organizational decision-making by addressing both strategic and operational considerations. To this end, the study poses the following Research Question (RQ):

RQ How can modern data management approaches (i.e., data platforms, data spaces, data mesh and data fabric) be delimited from each other?

Our study is organized as follows, including specific contributions to answering the question: First, the study outlines the development of data management, identifying areas that have not been fully explored through the review of existing literature. It then outlines the methodology of the systematic literature review, detailing the search approach and the process of gathering a comprehensive set of articles. In addition, the study presents visual representations of the distribution of significant literature and provides an in-depth examination of data platforms, data spaces, data meshes and data fabrics, including their similarities and differences. Finally, the study aims to guide future research by discussing the limitations, considering the theoretical and practical implications, and summarizing the insights gained from the analysis.

2 Theoretical background

Data management spans an extensive array of disciplines, encompassing a suite of processes essential for efficiently managing data across its entire lifecycle (K. Sharma et al. 2021). These processes are traditionally categorized into five core domains: collection, curation, consumption, conceptualization, and control (Chua et al. 2022). The evolution of data management within organizational contexts has transitioned from rudimentary administrative tasks and quality assurance practices to strategic frameworks that leverage big data as a driver of innovation and enhanced decision-making capabilities (Legner et al. 2020). This evolution is deeply intertwined with advancements in computational technologies, including the advent of big data analytics, cloud computing and the Internet of Things (IoT) (Chen et al. 2014). Historically, the transition from manual to electronic data processing during the mid-twentieth century marked a pivotal juncture in data management approaches (Gray 1996). The introduction of database management systems (DBMS) in the 1970s, particularly the relational model developed by Codd (1970), revolutionized data storage and retrieval methods, paving the way for modern relational databases and subsequent innovations in big data and cloud-based storage solutions (Mazumdar et al. 2019). These technological advancements have been instrumental in addressing the challenges posed by the increasing volume, velocity, and variety of data (Kashyap 2021), culminating in the sophisticated and interconnected

information systems observed today (Ghasemaghaei & Calic 2020; Oprel et al. 2024).

Modern data management faces several critical challenges, including the exponential growth of disparate and heterogeneous data, concerns over privacy, and the effective application of advanced analytics (Kumari & Muthulakshmi 2022). Central to addressing these challenges is the implementation of robust data governance frameworks and quality management protocols, which emphasize core principles such as accuracy, completeness, and consistency (Truong et al. 2020; Wang & Strong 1996). Distinctions in analytical paradigms further underscore the complexity of data management. Data warehouses, for instance, provide structured approaches optimized for analytical querying, whereas data lakes offer flexible storage solutions capable of accommodating raw, unstructured, and semi-structured data (Giebler et al. 2020). The increasing reliance on distributed approaches represents a fundamental shift in managing the complexity and scale of contemporary data systems (Paik et al. 2019). These approaches enhance scalability and flexibility by leveraging interconnected components that facilitate parallel processing and mitigate bottlenecks (Abadi 2009). Additionally, distributed approaches enable horizontal scalability, allowing for system growth through networked resource expansion rather than costly hardware upgrades (Stonebraker & Çetintemel 2018). The establishment of comprehensive metadata infrastructures is another cornerstone of modern data management, enabling organizations to address data quality issues effectively while navigating the complexities inherent in large-scale systems (Chen et al. 2014; Elouataoui et al. 2022). In the coming years, significant developments in the realm of data management are anticipated, characterized by the integration of artificial intelligence (AI) for automated governance, the evolution of adaptable and flexible data platforms, and the enhancement of privacy and security frameworks (Tian 2022). Decentralized architectures, such as data mesh, are emerging as critical innovations, emphasizing domain-specific data ownership and facilitating seamless data integration across diverse environments (Dehghani & Fowler 2022; Kuftinova, Maksimych, et al. 2022).

The existing literature provides a comprehensive examination of the theoretical underpinnings and practical implications of modern data management approaches (Machado et al. 2022a). Despite the substantial body of research on modern data management, notable gaps remain. Specifically, there is limited clarity on how these approaches address the operational complexities of diverse data landscapes or how their defining characteristics influence implementation effectiveness. These gaps underscore the necessity for further investigation into distinguishing and optimizing modern data management approaches—including data platforms, data mesh, data spaces, and data fabrics—for practical applications.

3 Research design

This study undertook a rigorous and expansive literature review to critically examine how the approaches "data platform", "data mesh", "data space" and "data fabric" are conceptualized, operationalized, and differentiated within the

field of Information Systems (IS) research. Employing a systematic research approach articulated by Webster and Watson (2002) and subsequently refined by vom Brocke et al. (2015), the review expanded its analytical scope to evaluate the implications of these approaches across academic, industrial, and governmental contexts. By situating data management approaches within their broader socio-technical landscapes, as emphasized by Benzies et al. (2006), the study integrated findings from both theoretical and practical research endeavors to construct a comprehensive understanding. The review utilized targeted keyword searches across multiple scholarly databases, including ACM, ScienceDirect, Springer, AIS eLibrary, Scopus, IEEE, and ProQuest. This effort culminated in the identification of an initial corpus comprising 1485 documents by December 2023, as visualized in Fig. 1.

The dataset underwent a thorough quality assessment, which involved removing 534 duplicate entries, leaving 951 unique sources for detailed evaluation. Each source was subjected to a meticulous full-text review to determine its relevance to the study's central research question. This review prioritized sources offering in-depth definitions, explanations, and critical analyses of the focal IS terms. Articles lacking substantive insights or merely referencing the terms in passing were excluded. For ambiguous cases, a detailed secondary review was conducted by the research team to uphold methodological rigor and minimize selection bias. Following the initial exclusion phase, reverse citation tracing was employed on key foundational works to identify additional relevant studies not indexed in the primary databases. This process adhered to the backward search methodology advocated by Kuhrmann et al. (2017). To further enhance the comprehensiveness of the dataset, supplementary searches were conducted using meta-search engines, capturing additional scholarly and non-indexed contributions. This iterative approach continued until theoretical saturation was achieved, ultimately adding 32 new articles to the dataset. The final corpus consisted of 189 articles: 157 derived from primary database searches and 32 sourced from supplementary search. Each article was annotated with detailed metadata to facilitate an in-depth analytical process.

The review approach, depicted in Fig. 1, was designed to prioritize transparency, reproducibility, and rigor, aligning with the principles of robust IS research as championed by vom Brocke et al. (2015). Two researchers independently conducted data extraction, focusing on delineating data management approaches, and collaboratively resolved any discrepancies. Table 1 presents an exemplary coding overview, providing a structured approach for categorizing the key characteristics of data management approaches while ensuring analytical depth and consistency.

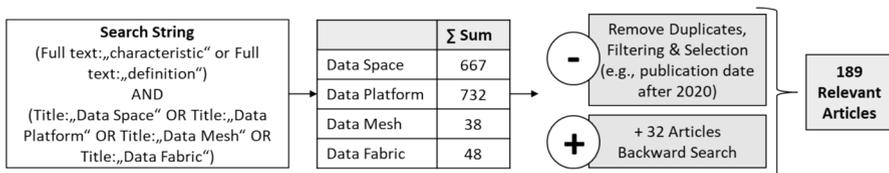


Fig. 1 Literature search process

Table 1 Example coding of emerging concepts

Quote from literature	Assigned code	Dimension
“While we have developed the approach and the data platform architecture implementation independently, the recently introduced concept called data mesh is also used to refer to a similar decentralised approach to data management.” (Loukiala et al. 2021, p. 2 f.)	Data mesh refers to a decentralized data management approach	Data mesh architecture
“Due to the progressive broadening of its definition and to the blurring of the architectural borderlines, the term DL is often replaced by the more general term data platform [5,6] or even data ecosystem [7,8].” (Francia et al. 2021, p. 299)	Data lakes are data platforms or data ecosystems	Data platform definition
“A key issue in this context is the respect of data sovereignty. The International Data Spaces (IDS) Reference Architecture has been developed in order to preserve data sovereignty.” (Nast et al. 2020, p. 1)	IDS reference architecture for data sovereignty	Data space architecture

The methodological rigor of this study enabled the creation of a nuanced approach for systematically categorizing and analyzing the defining characteristics of data management approaches. The process began with an abstract review to identify significant contributions and insights, which were subsequently integrated into broader academic discussions. This iterative approach facilitated the systematic organization of information, transitioning from general theoretical constructs to specific operational traits and applications. The final article data set list, included in the appendix, ensures full transparency, and supports the reproducibility and verification of the research findings.

4 Results

The following section contains the results of the systematic literature review. To this end, the authors have comprehensively evaluated all relevant publications. After a comprehensive overview of the literature, the authors define the terms data platform, data mesh, data space and data fabric in detail and show their origins and current applications. Based on the RQ, a summary of each term and its characteristic is provided in the text box at the end of each subsection. This is not an attempt to create a new definition, but rather to draw conclusions from the various descriptions of the articles in the literature review.

4.1 Descriptive findings

Research on data management first occurred in the mid of the twentieth century (Roach 1962). However, most publications, especially about data mesh and data fabric, stem from the last years. Figure 2 (left side) shows the cumulative number of publications for every approach in the last twenty years based on the Scopus database. Articles are counted in the statistics if they cover at least one of the four data

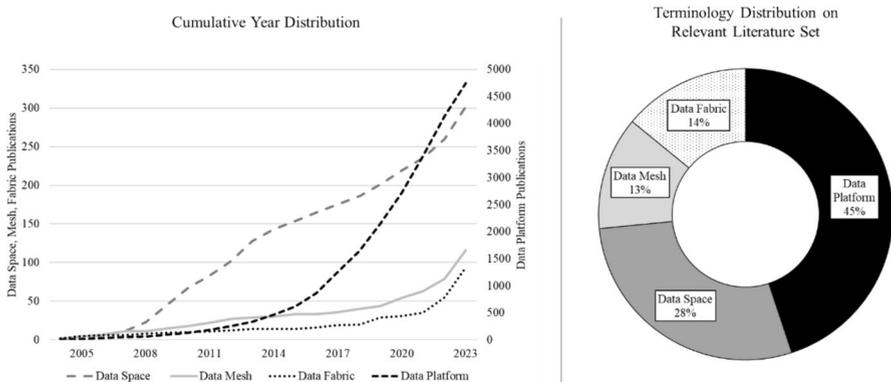


Fig. 2 Cumulative publication distribution of the past twenty years (left) and terminology distribution on the relevant literature set (right)

management approaches. Around 12,7% of the scanned literature refers to publications that do not exclusively focus on data platforms, data spaces, data meshes or data fabrics but observe more than one term. An example is the research of Hechler et al. (2023f) that mentions data fabric and data mesh. In this case, the authors counted this paper twice, one time for each category. The graph shows that data platform has always been the most discussed term, which is possibly due to many different potential interpretations of the phrase. However, the terms data space, data mesh and data fabric are also gaining increasing attention lately. More concretely, 44,9% of the analyzed 189 articles specialize in data platforms, while 28,5% focus on data spaces, 12,6% focus on data mesh and 14,0% cover data fabric, as summarized in Fig. 2 on the right side. The rapid rise in research requires a systematic review of previous findings to assist future research and avoid ambiguities in term definitions.

4.2 Data platform

The development and utilization of data platforms represent a cornerstone in the evolution of data management, transforming domains and enabling innovative applications across diverse domains (Borrero & Mariscal 2022). Data platforms form the foundation of modern digital ecosystems by facilitating data collection, processing, and monetization at an unprecedented scale (Beverungen et al. 2022). Unlike traditional digital platforms, data platforms emphasize the active exchange and utilization of data to foster value creation while addressing challenges related to governance or scalability (Bagheri et al. 2021; Reuver et al. 2022). They balance the need for openness with the risks of losing data sovereignty and privacy, offering specialized, context-dependent solutions that emphasize interoperability (Reuver et al. 2022). Hence, this evolution highlights the growing importance of governance approaches that range from centralized to decentralized control (Loukiala et al. 2021). The architectural evolution of data platforms has transitioned from traditional two-tier models, which separate transactional and analytical layers, to an lakehouse approach (Armbrust et al. 2021; Harby & Zulkernine 2022). The lakehouse approach combines the scalability of data lakes with the structured querying capabilities of data warehouses, enabling a unified approach to handle diverse workloads (Armbrust et al. 2021). Moreover, data platforms like the City Data and Analytics Platform (CiDAP) demonstrate how blockchain smart contracts and modular architectures can link data sources to smart city applications, fostering secure and efficient data sharing (Cheng et al. 2015; Nanthaamornphong et al. 2020).

Generally, data platforms serve as a transformative approach across various sectors (Fioriti et al. 2023). In urban environments, data platforms such as Snap4City provide scalable IoT-enabled infrastructures for real-time data streams, enabling applications in traffic management, environmental monitoring, and emergency response (Badii et al. 2018; Colarusso et al. 2022). Therefore, data platforms employ modular architectures that offer extensibility and advanced analytics, supporting sustainable and intelligent urban ecosystems (Cheng et al. 2015). By integrating data collection and analysis, they accelerate connected technology development for stakeholders (Lim et al. 2021). In agriculture, data

platforms address socio-economic challenges by empowering smallholder farmers with decision-support tools that integrate IoT and satellite data (Borrero & Mariscal 2022). Hence, data platforms emphasize transparency, data ownership, and equitable benefit-sharing while enabling predictive analytics for crop management and irrigation (Borrero & Mariscal 2022; Sedek et al. 2021). Domain-specific solutions such as these highlight the flexibility and specialization of data platforms required to address unique industry needs (Papageorgopoulos et al. 2022; Šestak & Copot 2023). Moreover, an open data platform for smart cities aimed at app developers may feature various components, including databases, semantically enriched databases, development kits, and reusable application components (Badii et al. 2018).

A recurring topic in data platform development is the trade-off between openness and control (Reuver et al. 2022). Open platforms encourage innovation and interoperability through APIs, modular extensions, and boundary resources, but simultaneously raise concerns over data misuse, privacy violations and ecosystem fragmentation (Beverungen et al. 2022; Reuver et al. 2022). These risks are exacerbated by market dynamics such as industry consolidation, which can lead to power asymmetries and disadvantage users in terms of pricing and contract terms (Turland & Slade 2020). Effective governance models, coupled with privacy-preserving technologies like multi-party computation, are critical to balancing these dynamics and maintaining trust (Rieyan et al. 2024; Šestak & Copot 2023). Privacy loss and data security are significant challenges, particularly under regulatory requirements such as the European Union's GDPR framework (Comandé & Schneider 2022). These regulations influence platform design and necessitate collaboration among diverse actors, including data owners, developers, and policymakers, to ensure sustainable governance and compliance (Rieyan et al. 2024).

Advanced management features, including metadata layers, transactional systems, and schema enforcement, enhance the reliability and quality of data platforms analytics (Ait Errami et al. 2023; Armbrust et al. 2021; Harby & Zulkernine 2022). These capabilities make platforms more resilient to inconsistencies while improving performance for machine learning workloads and analytics (Armbrust et al. 2021; Harby & Zulkernine 2022). Additionally, data platforms typically incorporate layered structures for raw, refined, public, and sensitive data, enabling secure transitions from production to analysis (Colarusso et al. 2022). The evolution of data platforms into broader ecosystems with diverse data analysis capabilities underscores their centrality in modern data management analysis (Francia et al. 2021). Data Platforms now integrate Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS) to enhance scalability and enable comprehensive services for data integration and security (Yao et al. 2021).

A **data platform** is a traditional comprehensive architecture, that facilitates storage, processing, and analysis of data within a unified platform. These platforms offer organizations the ability to manage their data lifecycle comprehensively from a single platform operator

4.3 Data space

Data spaces represent a novel approach in data management, emphasizing secure, interoperable, and sovereign data sharing across diverse organizations (Jarke et al. 2019). Unlike traditional centralized architectures, data spaces prioritize decentralization while maintaining interoperability through shared services and standardized protocols (Drees et al. 2021). This hybrid approach enables data owners to retain full control over their assets while engaging in heterogeneous ecosystems (Franklin et al. 2005). As organizations increasingly adopt data spaces, they address the complexities of data sovereignty, traceability, and compliance, establishing a foundation for trust in inter-organizational environments (Beverungen et al. 2022).

Data spaces are designed to manage and integrate heterogeneous data sources while guaranteeing their autonomy (Franklin et al. 2005). They emphasize coexistence rather than rigid integration, allowing incremental, pay-as-you-go scaling to address evolving organizational needs (Kotsev et al. 2020). A defining characteristic is data sovereignty, which ensures that data providers retain control over access and usage rights while enabling secure and compliant data sharing (Jarke et al. 2019). Interoperability is achieved through shared ontologies, semantic models, and standardized APIs, which enable seamless integration across systems (Curry et al. 2019). The modular nature of data spaces allows for incremental adoption and the addition of functionalities, ensuring scalability and adaptability (Brost et al. 2018). Furthermore, data spaces are uniquely capable of addressing dynamic interactions between datasets without requiring a centralized custodian, supporting diverse formats such as structured and unstructured data (Guo et al. 2021).

The architecture of data spaces integrate multiple components to ensure secure and efficient operations. Data providers and consumers form the foundational participants, responsible for generating and utilizing data within the ecosystem (Drees et al. 2021). These interactions are facilitated by standardized connectors (Gieß et al. 2024) and shared services that enforce secure and compliant exchanges (Bader et al. 2020). Data sovereignty frameworks allow participants to define and enforce ownership and usage policies, ensuring compliance with governance standards (Jarke et al. 2019). Semantic integration layers, such as knowledge graphs and ontologies, enhance discoverability and semantic alignment, allowing meaningful interactions across datasets (Curry et al. 2019). Federated governance approaches provide decentralized structures, ensuring that shared rules and standards are adhered to while maintaining local autonomy (Hutterer & Krumay 2022). Additionally, the integration of International Data Spaces (IDS) standards enables secure, trusted, and scalable data exchange, further enhancing interoperability and governance (Brost et al. 2018).

Data sovereignty ensures that providers maintain full control over their data by defining access and usage rights (Bader et al. 2020). Interoperability is supported through standardized metadata, shared semantic frameworks, and APIs, which enable seamless communication between systems (Hutterer et al. 2023). Additionally, data spaces integrates heterogeneous data and supports intelligent decision-making without a centralized database, evolving from early relational databases to systems with a flexible, pay-as-you-go approach (Guo et al. 2021). Interoperability enhances

scalability, allowing for incremental integration of participants and functionalities as the data space evolves (Kotsev et al. 2020). Robust mechanisms for trust and security, such as encryption, identity management, and federated governance frameworks, ensure compliance and foster confidence among participants (Bader et al. 2020; Jarke et al. 2019). Privacy-preserving mechanisms like secure communication and trust management further enhance usability and regulatory compliance (Alonso et al. 2018). Feature such as best-effort query capabilities are also a key aspects of data space design, enabling continuous operational functionality even in complex environments (Franklin et al. 2005).

Data spaces have transformative applications across industries, showcasing their versatility and impact (Gieß et al. 2025). For example, in Industry 4.0 and smart manufacturing, data spaces facilitate real-time data sharing and predictive analytics, optimizing production processes and enabling autonomous decision-making (Alonso et al. 2018). These capabilities enhance operational efficiency and reduce costs (Alonso et al. 2018). In smart cities, integrated IoT data streams enhance urban planning, traffic optimization, and citizen services, enabling sustainable urban management (Curry et al. 2019). The energy management sector benefit from real-time data integration, which supports optimized resource allocation and sustainability initiatives (Berkhout et al. 2022).

Future directions for data spaces include the integration of advanced technologies such as AI, blockchain, and edge computing (Brée et al. 2024). AI and machine learning enhance data analytics and predictive capabilities, unlocking new insights and efficiencies (Tardieu 2022). Blockchain introduces decentralized trust mechanisms, ensuring secure and transparent data sharing (Pampus et al. 2022). Edge computing supports real-time processing at the network edge, enhancing decision-making in distributed environments (Gallay et al. 2017). Initiatives such as GAIA-X aim to create globally compliant and interoperable data spaces, addressing challenges in privacy, security, and governance while fostering collaboration (Tardieu 2022). Furthermore, data spaces are expected to evolve into foundational applications for public–private collaboration, enabling a wide range of data-driven services and applications (Drees et al. 2021). By emphasizing data sovereignty, interoperability, and trust, they empower stakeholders to collaborate securely while maintaining control over their data.

A **data space** is a federated architecture that refers to a decentralized data infrastructure where organizations can securely exchange data while maintaining sovereignty over their data assets. It supports collaboration by enabling controlled access to data across organizational boundaries

4.4 Data mesh

The data mesh paradigm represents a transformative approach to organizational data management, addressing the limitations of traditional centralized approaches such as data lakes and warehouses (Bode et al. 2023; Machado et al. 2022b). Emerging architectures such as the data mesh approach further decentralize control, empowering domains to manage their data autonomously and independently (Loukiala et al.

2021). By decentralizing data ownership and aligning it with domain expertise, data mesh fosters scalability and adaptability within organizations (Dahdal et al. 2023). Unlike monolithic architectures, this approach treats data as a product, emphasizing its usability, quality and lifecycle (Pongpech 2023). As a socio-technical approach, data mesh enables organizations to become more data-driven, empowering individual domains to manage their data autonomously while adhering to governance requirements (Machado et al. 2022b).

Data mesh is built on four foundational principles: domain-oriented data ownership, data as a product, self-serve data infrastructure, and federated governance (Machado et al. 2022a). Each domain is a self-contained unit, responsible for managing its data throughout its lifecycle, thereby minimizing bottlenecks caused by centralized IT teams (Loukiala et al. 2021). Architecturally, data mesh comprises several key components. Data domains are foundational, with each domain managing its own data lifecycle independently (Machado et al. 2022a). These domains are supported by a self-serve infrastructure that enables autonomous data processing and sharing without reliance on centralized systems (Joshi et al. 2021). This decentralization allows those closest to the data source to take full responsibility for its quality and usability (Joshi et al. 2021). The principle of "data as a product" ensures that data is treated with the same rigor as consumer products, emphasizing discoverability, reliability, and user-focused design (Wrembel 2023). Self-serve infrastructure equips domains with tools and platforms to autonomously create, maintain, and share data products (Pongpech 2023). Federated governance establishes decentralized yet standardized policies for compliance, security, and interoperability across the organization (Dahdal et al. 2023).

Governance frameworks ensure interoperability and compliance through standardized rules, ensuring that the decentralized approach operates cohesively across the organization (Wider et al. 2023). Critical elements such as APIs, metadata layers, and data catalogs enhance interoperability and streamline data discovery and integration across domains (Machado et al. 2022a). In addition, the core concept of the architecture is the data product, which interacts with the mesh via ports, supported by an operator app for orchestrating tools and managing product lifecycles, a data catalog for enhancing discoverability and lineage, and a governance layer for implementing rules like access control, all integrated to provide APIs for both the data product developer experience plane and the mesh experience plane (Wider et al. 2023). The concept aligns with DATSIS principles—discoverability, addressability, trustworthiness, self-description, interoperability, and security—to ensure data quality and usability (Dahdal et al. 2023).

The core concept of the data mesh architecture is the data product, which interacts with the mesh through well-defined interfaces, supported by governance frameworks and tools for lifecycle management (Machado et al. 2022a). Operational components like mesh catalogs, metadata layers, and communication channels facilitate efficient development and management of domain-aligned data products (Machado et al. 2022b). Data stewardship is decentralized, with each data source maintaining integrity and accessibility through uniform interfaces, enabling robust governance and interoperability (Wrembel 2023). By decentralizing data ownership and emphasizing data as a product, it addresses the inefficiencies of centralized systems

and enables organizations to scale effectively (Machado et al. 2022b). The data mesh approach shifts the focus from technology to data as the core of value creation, emphasizing quality, transparency, and responsibility through data domains, reliable access enforced by governance rules, and enhanced interoperability through standards (Krystek et al. 2023). Consequently, data product advertisements should detail traits like discoverability for easy web indexing, clarity and explicitness in descriptions, manageable direct references for workflow integration, compliance with standards in data handling, robust security measures for access control, verified quality for trustworthiness, and valuation to highlight potential business profits and applicability (Krystek et al. 2023).

A **data mesh** is a decentralized architectural approach that treats data as a product. Ownership and accountability for data quality and governance are distributed across different domains (i.e., cross-functional teams)

4.5 Data fabric

Data fabric represents an advanced data management paradigm that leverages AI and automation to unify access to diverse data sources, integrating them seamlessly across decentralized architectures while maintaining data consistency and governance (Strengtholt 2023). It integrates and manages structured and unstructured data across on-premises, cloud, and edge environments, providing seamless real-time access without centralizing data into a single repository (Hechler et al. 2023d). Unlike traditional data integration approaches, data fabric employs metadata-driven methodologies and semantic knowledge graphs to optimize data handling and reduce silos, thereby enhancing operational efficiency and scalability (Hechler et al. 2023b; Li et al. 2022). It acts as the foundational layer, bridging gaps between disparate systems and promoting a decentralized yet unified data management approach (Hechler et al. 2023f). By focusing on real-time analytics and operational efficiency, data fabric enables organizations to address challenges such as data silos and regulatory compliance (Hechler et al. 2023c). This is achieved through its robust support for scalable architectures, advanced encryption mechanisms, and adherence to frameworks like GDPR and HIPAA (Rieyan et al. 2024; Saini et al. 2023). Unlike data mesh, which prioritizes domain-centric data product creation, data fabric provides the necessary technical infrastructure to support such decentralized systems, creating a symbiotic relationship between the two approaches (Hechler et al. 2023e).

The architecture of data fabric is typically organized into a modular, multi-layer framework to manage complex data environments. Several studies propose five-layer architectures, including collection, storage, processing, analytics, and application layers (Kuftinova et al. 2023; Zhu et al. 2019). These frameworks aim to integrate data across various platforms and storage locations, enabling seamless access and sharing in distributed environments (Anand 2022; Kuftinova et al. 2023). Some researchers suggest alternative layer configurations, such as four-layer models (Nema 2010; Zhao et al. 2016) or domain-oriented decomposition approaches (Kuftinova et al. 2023). Data fabric architectures can be designed as modular microservices or

integral single fabric systems, with their effectiveness varying across industries (Cao & Iansiti 2022). Implementing data fabric requires careful consideration of domain models and data modeling approaches (Östberg et al. 2022), as well as the use of specialized tools and languages for data processing (Kuftinova, Maksimychev, et al. 2022). Data fabric distinguishes itself by consolidating organizational knowledge into a unified schema through ontological modeling (Kuftinova, Maksimychev, et al. 2022). This integration is further enhanced by supporting technologies such as machine learning, cloud-based solutions, and microservices, which streamline data operations and minimize inefficiencies (Kuftinova, Maksimychev, et al., 2022).

The design of data fabric revolves around four primary principles: modularity, interoperability, scalability, and security (Strengtholt 2023). Interoperability is achieved through shared ontologies, semantic frameworks, and standardized APIs, ensuring seamless data integration across heterogeneous environments (Hechler et al. 2023e). Scalability is a core feature, allowing organizations to expand their data ecosystems without compromising performance or security (Hechler et al. 2023c; Strengtholt 2023). Meanwhile, privacy and security are reinforced by advanced techniques like partially homomorphic encryption and federated learning, which protect sensitive data while enabling collaborative analytics (Saini et al. 2023). Notable features include real-time data sharing, best-effort query capabilities, and robust regulatory compliance mechanisms, which enhance the usability and governance of data fabric architectures (Rieyan et al. 2024).

Data fabric has transformative applications across various industries. For instance, its federated learning capabilities allow institutions to analyze medical images collaboratively while adhering to privacy regulations (Theodorou et al. 2021). In manufacturing and logistics, data fabric unifies supply chain data, facilitating predictive maintenance and enhancing operational efficiency in Industry 4.0 environments (Kuftinova, Ostroukh, et al. 2022a, b). Financial services leverage its real-time analytics capabilities to optimize portfolio management and enhance fraud detection (Hechler et al. 2023f). Looking to the future, platforms like the MEDAL architecture exemplify the potential for runtime adaptation and cloud-native data fibers, driving automation and enabling intelligent, data-driven decisions (Theodorou et al. 2021). Data fabric represents a paradigm shift in data management, offering a flexible, scalable, and secure approach to unify disparate data sources across complex ecosystems (Saini et al. 2023). By leveraging AI, metadata-driven methodologies, and semantic knowledge graphs, it addresses critical challenges in data governance, interoperability, and scalability (Li et al. 2022).

A **data fabric** is an emerging centralized integration approach that focuses on creating a seamless data environment across different business units within an organization, allowing for integrated data management and governance.

4.6 Delimitation

The findings from this comprehensive review illuminate critical dimensions of how the four focal IS terms are conceptualized, applied, and differentiated across diverse

contexts. These distinctions are pivotal for advancing both theoretical understanding and practical applications within the realm of data management approaches, particularly in applications requiring real-time data integration and advanced analytics, such as smart cities and IoT ecosystems. Data platforms emerged as foundational systems, typically characterized by centralized architectures that enable robust operational analytics and streamlined data workflows. These setting-oriented data platforms serve as the flexible cornerstone for many organizations, particularly those requiring centralized solutions to manage large-scale operational and transactional analytics. Their emphasis on structured data workflows and centralized control optimizes performance for standardized use cases, making them essential for organizations with significant data processing needs. In contrast, data meshes have gained prominence as scalable solutions for large, multifaceted organizations. These domain-oriented architectures promote decentralization by embedding data ownership and responsibility within specific organizational domains. By empowering individual domains to manage their data as independent products, data meshes enhance scalability and agility, facilitating rapid innovation while maintaining consistency across broader organizational settings. Data spaces were conceptualized as a decentralized data infrastructure emphasizing governance, sovereignty, and interoperability. They enable multi-stakeholder data sharing while maintaining individual control, fostering data availability through enhanced interoperability and seamless data exchanges across organizational boundaries. This approach is particularly critical in sectors such as healthcare and manufacturing, where data sovereignty and regulatory compliance are paramount. Meanwhile, data fabrics were defined as integration-centric solutions leveraging advanced technologies like AI for automation to unify disparate data sources and facilitate seamless data flows across complex data ecosystems. These holistic, technology-driven approaches dynamically adapt to evolving data landscapes, ensuring efficient data integration and accessibility in diverse organizational contexts as a further extension of data platforms capabilities. As the field continues to evolve, these findings underscore the importance of adaptive, context-sensitive approaches to data management that align with the unique demands of diverse organizational and societal contexts. Table 2 provides a comparative analysis of the four data management approaches.

Each approach is examined through key aspects, including architecture, characteristics, governance, goal, scope, key strengths, use case and examples. By outlining these aspects, Table 2 aims to clarify the strengths and applications of each approach, offering conceptual clarity to assess their suitability for various organizational applications.

Our study emphasizes the transition from centralized to decentralized approaches. Centralized approaches, such as *data platforms* and *data fabrics*, provide a unified framework for data management, improving accessibility and governance. By consolidating data into a common infrastructure, e.g., within an organization, these systems streamline operations and strengthen governance control. However, they can sacrifice flexibility, responding slowly to changing requirements. In contrast, decentralized approaches like *data meshes* and *data spaces* prioritize agility and autonomy through federated governance, promoting secure data sharing without a central custodian. This allows for a dynamic data

Table 2 Delimitation of data management approaches

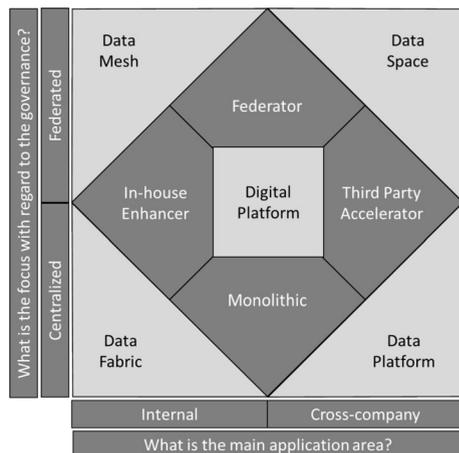
Approach	Data platform	Data space	Data mesh	Data fabric
Illustration				
Characteristics	Consolidated data in a single location, schema-on-write	Data remains with owners, standardized protocols for access	Each domain manages its data independently	AI-powered analysis, integration, and real-time data delivery
Governance	Centralized	Federated, based on common standards	Federated, domain-specific	Automated, AI-supported
Goal	Unified view of data	Secure and sovereign data exchange	Flexibility for decentralized teams	Seamless integration of heterogeneous data sources
Scope	Centralized ecosystem for analytics and operations	Open and governed ecosystem for secure data sharing	Flexible architecture for domain-specific needs	Dynamic integration of heterogeneous data environments
Key strengths	Comprehensive data lifecycle management	Sovereign data exchange	Decentralized data ownership	AI-driven integration
Use case	Data analysis and management	Data sharing between organizations	Provision of data products	Integration of heterogeneous data sources
Example	Amazon Redshift (2025), Snowflake (2025)	SCSN (2023), Catena-X (2023)	Nefflix Data Mesh (2022)	IBM Data Fabric (2025)
Example literature	Bartelheimer et al. (2022), Francia et al. (2021)	Möller et al. (2024), Gieß et al. (2025)	Dehghani and Fowler (2022), Bode et al. (2023)	Hechler et al. (2023a), Alvord et al. (2022)
Legend	Data intermediary ●	Data user○	Data provider ●	Organization

infrastructure that quickly adapts to new opportunities, such as increased data sovereignty, or evolving needs, like interoperability in the digital era. The delineation of insights into data management approaches reflects various philosophies around federalization, governance, and sovereignty. Centralized models, such as data platforms and data fabrics, focus on creating a cohesive platform, while decentralized approaches, like data meshes and data spaces, emphasize flexibility and decentralized data control. Our data management matrix, outlined in Fig. 3, identifies four key roles: Federator, Third Party Accelerator, Monolithic, and In-house Enhancer. This framework helps organizations select and adapt the most suitable data management approach for their specific needs within data sharing contexts. It allows them to choose between a centralized structure that prioritizes control and governance, and a decentralized one that offers agility and rapid response to changing demands.

To navigate the increasingly complex landscape of data management, organizations must consider two primary dimensions: governance mode and application area. Governance mode delineates centralized approaches, which offer unified governance mechanisms, from decentralized approaches that distribute governance across domains. Application area differentiates approaches intended for in-house data sharing within a single organization from those facilitating inter-organizational data sharing (third-party). This classification informs the structured data management matrix, which comprises four primary roles:

- **Monolithic Role:** Centralizes data management, enabling direct access to data and related services through platforms operated by a single entity. These systems emphasize control and efficiency.
- **Federator Role:** Supports a broader data-sharing ecosystem by offering services such as identity verification and metadata cataloging without directly managing data. Federators establish trust and interoperability within multi-entity networks.
- **Third Party Accelerator Role:** Optimizes cross-organizational data sharing by streamlining data flows, enhancing interoperability, and enabling collaborative

Fig. 3 Data management matrix



data usage. Accelerators are critical for industries with extensive third-party interactions.

- **In-house Enhancer Role:** Refines internal data management processes through advanced integration, analytics, and governance frameworks, driving operational excellence within organizations.

These roles can operate independently or in conjunction with one another. For example, a Monolithic system may oversee a centralized data platform with comprehensive control, while a Federator might facilitate mediation platforms that support proprietary and third-party services without direct data oversight. This flexibility allows organizations to tailor their data management approach to align with specific operational needs and existing digital infrastructures.

In-house approaches, such as data meshes and data fabric, often serve as Enhancers, streamlining internal workflows and improving governance. Third-party approaches, like data platforms and data spaces, typically act as Accelerators, facilitating broader cross-organizational data sharing. While these roles are conceptually distinct, their practical implementation frequently involves hybrid configurations. For instance, a data fabric may integrate with a data mesh to combine the benefits of centralized integration with decentralized autonomy. Similarly, interconnected data spaces can form a comprehensive data space mesh network, delivering scalable and interoperable solutions (Drees et al. 2022). Such integrations illustrate the flexibility and adaptability of modern data management approaches, enabling organizations to effectively respond effectively to the complexities of contemporary environments.

The conceptual data management matrix not only provides a robust framework for categorizing and understanding these approaches but also highlights their practical applications and benefits. It acknowledges that while centralized and decentralized approaches offer distinct advantages, their interplay often results in innovative hybrid solutions tailored to organizational requirements.

By leveraging this matrix, organizations can strategically evaluate and implement data management approaches that align with their operational, strategic, and regulatory objectives. Such a structured approach ensures resilience and efficiency in the face of continual technological advancements and organizational changes. Furthermore, this adaptability positions organizations to harness the full potential of their data assets, enhancing decision-making and maintaining a competitive edge.

5 Discussion

This study undertakes a comprehensive and critical examination of contemporary data management approaches, including data platforms, data spaces, data meshes, and data fabrics, to provide a robust theoretical foundation and actionable insights for advancing the field. By evaluating their similarities, differences, and contextual applications, this research addresses pivotal challenges in data sharing and management. Centralized approaches, such as data platforms and data fabrics, are characterized by unified governance, operational consistency, and enhanced security measures. These approaches are particularly advantageous for organizations

with substantial and complex data needs. However, their inherent rigidity often limits their ability to adapt to dynamic and rapidly evolving business environments. In contrast, decentralized approaches like data meshes and data spaces prioritize agility, autonomy, and federated governance. These approaches enable secure data sharing and collaboration, reflecting a growing demand for approaches that balance governance and control with the adaptability required to meet evolving technological and business requirements.

Although prior research has clarified distinctions among data warehouses, data lakes, data lakehouses, data meshes, and data fabrics (Bode et al. 2023), significant gaps remain in understanding the nuanced differences between data platforms, data spaces, and broader data ecosystems (Beverungen et al. 2022). Recent studies emphasize the critical role of flexible data management approaches to enable federated data sharing (Hechler et al. 2023d). Among these, data spaces have gained traction for their ability to address pressing issues such as data sovereignty and interoperability (Scherenberg et al. 2024). Political initiatives, such as those led by the European Commission, underscore the importance of data spaces in fostering reliable and sovereign data ecosystems (European Commission 2023). In this context, data spaces are increasingly recognized as essential approach for maintaining data sovereignty within multi-organizational environments (Hutterer & Krumay 2022).

Data fabrics and data meshes represent distinct yet complementary roles in modern data management approaches. Data fabrics excel in constructing AI-driven architectures to integrate and manage diverse data sources. By leveraging advanced analytics and automation, data fabric streamlines workflows and provides centralized solutions for intricate data integration challenges. Conversely, data meshes decentralize governance by empowering domain-specific teams to create and manage self-service data products. This paradigm fosters organizational autonomy, agility, and localized decision-making (Hechler et al. 2023a). By systematically reviewing both academic and industry literature, this study elucidates the roles and contributions of these approaches, emphasizing their increasing relevance since 2019 (see Fig. 2). The thematic classification outlined in Fig. 3 extends these insights (Bode et al. 2023), offering significant implications for advancing Information Systems (IS) research and related disciplines.

The study's findings demonstrate that each data management approach is tailored to specific scenarios, with distinct implications for practical applications. Data platforms serve as an approach for cross-organizational data sharing, acting as centralized contact points for stakeholders. CiDAP, for instance, integrates multiple data sources, offering a unified platform for managing smart city data (Cheng et al. 2015). Data fabrics are particularly effective in automating data management by incorporating advanced technologies, including AI. They are ideally suited for centralizing data workflows and integrating disparate data sources within organizations. For example, the M-Data-Fabric architecture employs metadata-driven techniques to unify and analyze enterprise data (Liu et al. 2022), while the MEDAL Data Fabric automates data orchestration across data and resource layers (Theodorou et al. 2021). Decentralized approaches, such as data meshes, focus on federated governance and domain-specific autonomy to meet internal organizational needs. For example, a data mesh implemented in smart city applications enhances operational

efficiency and governance structures (Krystek et al. 2023). Similarly, data spaces enable trusted, cross-organizational data sharing through secure and reliable data infrastructures, as exemplified by the Mobility Data Space (Drees et al. 2021).

Collaboration and interoperability among these paradigms are crucial to prevent conceptual fragmentation and ensure cohesive digital ecosystems. Initiatives like the International Data Spaces (IDS) and GAIA-X directly address these challenges by promoting principles of data sovereignty and fostering trust (Tardieu 2022). Additionally, the concept of a data space mesh, which interconnects multiple data spaces, offers a scalable and adaptable framework for managing data within trustworthy ecosystems (Drees et al. 2022). Data fabrics further enhance these initiatives by providing the technical foundation for data meshes while retaining centralized oversight when necessary (Hechler et al. 2023a). By understanding the unique strengths and limitations of each paradigm, organizations can develop tailored approaches that optimize data management practices. Ultimately, the integration and collaboration of these approaches will be instrumental in advancing data ecosystems, ensuring their adaptability and resilience amid ongoing technological and organizational transformation.

6 Conclusion, limitations, and future research

This study provides a detailed exploration of contemporary data management approaches, specifically data platforms, data spaces, data meshes, and data fabrics. By elucidating their distinctions, the study demonstrates that each approach offers unique benefits tailored to the application contexts, addressing diverse organizational needs and data management challenges. Data platforms and data fabrics, with their more centralized structure, provide a cohesive approach for managing and controlling data, ensuring broad accessibility and streamlined management. Data meshes and data spaces, on the other hand, advocate decentralization and focus on domain-specific data ownership or federated platform ecosystems that improve data sovereignty for cross-organizational collaboration.

While this study offers significant contributions, it also acknowledges several limitations. A primary challenge lies in the lack of empirical evidence directly comparing the real-world effectiveness of these paradigms, as research in this area remains nascent. The broad scope of this study, while comprehensive, may not capture all possible organizational scenarios or industry-specific requirements, potentially limiting the applicability of its findings. Additionally, the dataset analyzed reflects a snapshot of the continuously expanding body of literature on data management, underscoring the need for ongoing updates to ensure relevance and comprehensiveness.

Future research should address these limitations through rigorous empirical investigations that evaluate the adoption, implementation, and operational impact of these approaches across diverse contexts. Moreover, exploring the interoperability of these paradigms within existing digital infrastructures would yield valuable guidance for practitioners. Examining their adaptability to evolving data governance requirements, driven by technological advancements and regulatory shifts,

could further refine their practical utility and broaden their applicability. An especially promising area for future exploration involves the development of hybrid approaches that integrate the strengths of centralized and decentralized approaches. Hybrid approaches could leverage the governance and integration efficiencies of centralized approaches while incorporating the flexibility, scalability, and agility of decentralized ones. These approaches have the potential to address a wider array of organizational demands, ranging from enhancing data sovereignty to supporting comprehensive digital transformation initiatives. Investigating the optimal configurations and practical implementations of such hybrid solutions would contribute to the evolution of resilient and versatile data management approaches. For instance, studying the integration of data fabrics with data meshes within specific industries could provide insights into their combined potential to deliver scalable and adaptive solutions. Similarly, examining the role of emerging technologies, such as AI and machine learning, in enhancing these paradigms would offer valuable perspectives on their future trajectories.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10257-025-00707-4>.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abadi DJ (2009) Data management in the cloud: limitations and opportunities. *IEEE Data Eng Bull* 32(1):3–12
- Achanta A, Boina R (2023) Evolving paradigms of data engineering in the modern era: challenges, innovations, and strategies. *Int J Sci Res (IJSR)* 12(10):606–610
- Ait Errami S, Hajji H, Ait El Kadi K, Badir H (2023) Spatial big data architecture: from data warehouses and data lakes to the LakeHouse. *J Parallel Distr Comput* 176:70–79. <https://doi.org/10.1016/j.jpdc.2023.02.007>
- Alonso Á, Pozo A, Cantera J, de La Vega F, Hierro J (2018) Industrial data space architecture implementation using FIWARE. *Sensors* 18(7):2226. <https://doi.org/10.3390/s18072226>
- Alvord MM, Lu F, Du B, Chen CA (2020) Big data fabric architecture: How big data and data management frameworks converge to bring a new generation of competitive advantage for enterprises. *Enterprise Arch Prof J*
- Anand JV (2022) Digital transformation by data fabric. *JSWS* 4(3):196–201. <https://doi.org/10.36548/jsws.2022.3.006>
- Armbrust M, Ghodsi A, Xin R, Zaharia M (2021) Lakehouse: a new generation of open platforms that unify data warehousing and advanced analytics. In: proceedings of CIDR.

- Azad P, Navimipour NJ, Rahmani AM, Sharifi A (2020) The role of structured and unstructured data managing mechanisms in the Internet of things. *Clust Comput* 23(2):1185–1198. <https://doi.org/10.1007/s10586-019-02986-2>
- Bader S, Pullmann J, Mader C, Tramp S, Quix C, Müller AW, Akyürek H, Böckmann M, Imbusch BT, Lipp J, Geisler S, Lange C (2020) The international data spaces information model – an ontology for sovereign exchange of digital content. In: Pan JZ, Tamma V, d’Amato C, Janowicz K, Fu B, Polleres A, Seneviratne O, Kagal L (eds) *The Semantic Web – ISWC 2020*. Springer International Publishing, Cham
- Badii C, Belay EG, Bellini P, Cenni D, Marazzini M, Mesiti M, Nesi P, Pantaleo G, Paolucci M, Valtolina S, Soderi M, Zaza I. (2018) Snap4City: A scalable IOT/IOE platform for developing smart city applications. In: 2018 IEEE smartworld, ubiquitous intelligence & computing, advanced & trusted computing, scalable computing & communications, cloud & big data computing, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI). IEEE
- Bagheri S, Brandt T, Sheombar HS, Van Oosterhout M (2021) Value Creation through Urban Data Platforms: A Conceptual Framework. In: *Proceedings of the 54th Hawaii International Conference on System Sciences*, Hawaii, USA.
- Bartelheimer C, Zur Heiden P, Lüttenberg H, Beverungen D (2022) Systematizing the lexicon of platforms in information systems a data-driven study. *Electron Markets*. 32(1):375–396. <https://doi.org/10.1007/s12525-022-00530-6>
- Benzies KM, Premji S, Hayden KA, Serrett K (2006) State-of-the-evidence reviews: advantages and challenges of including grey literature. *Worldviews Ev Based Nurs* 3(2):55–61. <https://doi.org/10.1111/j.1741-6787.2006.00051.x>
- Bergman K, Wang H (2013) Optically interconnected high performance data centers. In: Kachris C, Bergman K, Tomkos I (eds) *Optical interconnects for future data center networks*. Springer, New York, USA
- Beverungen D, Hess T, Köster A, Lehrer C (2022) From private digital platforms to public data spaces: Implications for the digital transformation. *Electron Mark* 32(2):493–501. <https://doi.org/10.1007/s12525-022-00553-z>
- Bode J, Kühl N, Kreuzberger D, Hirschl S, Holtmann C (2023) Towards avoiding the data mess: industry insights from data mesh implementations. *Adv Online Pub*. <https://doi.org/10.48550/ARXIV.2302.01713>
- Borrero JD, Mariscal J (2022) A case study of a digital data platform for the agricultural sector: a valuable decision support system for small farmers. *Agriculture* 12(6):767. <https://doi.org/10.3390/agriculture12060767>
- Brée T, Karger E, Ahlemann F (2024) Shaping the future of data ecosystem research—what is still missing? *IEEE Access* 12:103162–103175. <https://doi.org/10.1109/ACCESS.2024.3432969>
- Brost GS, Huber M, Weiß M, Protsenko M, Schütte J, Wessel S (2018) An ecosystem and IoT Device architecture for building trust in the industrial data space. In: *proceedings of the 4th ACM workshop on cyber-physical system security*. <https://doi.org/10.1145/3198458.3198459>
- Cao R, Iansiti M (2022) Cloud adoption and digital transformation: the paradoxical role of enterprise data architecture. *Advance online publication, SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4307486>
- Catena-X. (2023). *Automotive Network*. <https://catena-x.net/en/>
- Chen M, Mao S, Liu Y (2014) Big data: a survey. *Mobile Netw Appl* 19(2):171–209. <https://doi.org/10.1007/s11036-013-0489-0>
- Cheng B, Longo S, Cirillo F, Bauer M, Kovacs E (2015) Building a big data platform for smart cities: experience and lessons from santander. In: 2015 IEEE international congress on big data (pp. 592–599). IEEE. <https://doi.org/10.1109/BigDataCongress.2015.91>
- Chua C, Indulska M, Lukyanenko R, Maass W, Storey VC (2022) MISQ research curation on data management. *MIS Quarterly*
- Codd EF (1970) A relational model of data for large shared data banks. *Commun ACM* 13(6):377–387. <https://doi.org/10.1145/362384.362685>
- Colarusso C, De Iasio A, Furno A, Goglia L, Merzoug MA, Zimeo E (2022) Promenade: a big data platform for handling city complex networks with dynamic graphs. *Future Generat Comput Syst* 1(137):129–145. <https://doi.org/10.2139/ssrn.4014899>
- Comandé G, Schneider G (2022) It’s time: Leveraging the GDPR to shift the balance towards research-friendly EU data spaces. *Common Market Law Rev* 59(3):739–776. <https://doi.org/10.54648/COLA2022051>

- European Commission (2023) European Data Strategy: Making the EU a role model for a society empowered by data. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en
- Curry E, Derguech W, Hasan S, Kouroupetroglou C, ul Hassan U (2019) A real-time linked dataspace for the internet of things: enabling “pay-as-you-go” data management in smart environments. *Future Generat Comput Syst* 90:405–422
- Dahdal S, Poltronieri F, Tortonesi M, Stefanelli C, Suri N (2023) A Data Mesh Approach for Enabling Data-Centric Applications at the Tactical Edge. In: 2023 international conference on military communications and information systems (ICMCIS). IEEE. <https://doi.org/10.1109/ICMCIS59922.2023.10253568>
- Dehghani Z, Fowler M (2022) *Data mesh: delivering data-driven value at Scale* (First edition). O’Reilly.
- Drees H, Kubitzka DO, Lipp J, Pretzsch S, Langdon CS (2021) Mobility data space—first implementation and business opportunities. In: ITS World Congress
- Drees H, Pretzsch S, Heinke B, Wang D, Schlueter Langdon C (2022) Data space mesh: interoperability of mobility data Spaces. In 14th ITS European congress
- Elouataoui W, El Alaoui I, Gahi Y (2022) Metadata quality in the era of big data and unstructured content. In: Maleh Y, Alazab M, Gherabi N, Tawalbeh L, Abd El-Latif AA (eds) *Lecture notes in networks and systems advances in information, communication and cybersecurity*. Springer International Publishing, Cham
- Fioriti D, Stevanato N, Ducange P, Marcelloni F, Colombo E, Poli D (2023) Data platform guidelines and prototype for microgrids and energy access: matching demand profiles and socio-economic data to foster project development. *IEEE Access* 11:73218–73234. <https://doi.org/10.1109/ACCESS.2023.3294841>
- Francia M, Gallinucci E, Golfarelli M, Leoni AG, Rizzi S, Santolini N (2021) Making data platforms smarter with MOSES. *Futur Gener Comput Syst* 125:299–313. <https://doi.org/10.1016/j.future.2021.06.031>
- Franklin M, Halevy A, Maier D (2005) From databases to dataspace: a new abstraction for information management. *ACM SIGMOD Rec* 34(4):27–33. <https://doi.org/10.1145/1107499.1107502>
- Gallay O, Korpela K, Tapio N, Nurminen JK (2017) A peer-to-peer platform for decentralized logistics. In: proceedings of the hamburg international conference of logistics (HICL)
- Geisler S, Vidal M-E, Cappiello C, Lóscio BF, Gal A, Jarke M, Lenzerini M, Missier P, Otto B, Paja E, Pernici B, Rehof J (2022) Knowledge-driven data ecosystems toward data transparency. *J Data and Inf Quality* 14(1):1–12. <https://doi.org/10.1145/3467022>
- Gelhaar J, Müller P, Bergmann N, Dogan R (2023) Motives and incentives for data sharing in industrial data ecosystems: an explorative single case study. In: proceedings of the 56th hawaii international conference on system sciences (HICSS). <https://hdl.handle.net/10125/103085>
- Ghasemaghaei M, Calic G (2020) Assessing the impact of big data on firm innovation performance: big data is not always better data. *J Bus Res* 108:147–162. <https://doi.org/10.1016/j.jbusres.2019.09.062>
- Giebler C, Gröger C, Hoos E, Schwarz H, Mitschang B (2019) Leveraging the data lake: current state and challenges. In: Ordonez C, Song I-Y, Anderst-Kotsis G, Tjoa AM, Khalil I (eds) *Big Data Analytics and Knowledge Discovery*. Springer International Publishing, Cham
- Giebler C, Gröger C, Hoos E, Schwarz H, Mitschang B (2020) A zone reference model for enterprise-grade data lake management. In: 2020 IEEE 24th international enterprise distributed object computing conference (EDOC) (pp. 57–66). IEEE. <https://doi.org/10.1109/EDOC49727.2020.00017>
- Gieß A, Schoormann T, Möller F, Gür I (2025) Discovering data spaces: a classification of design options. *Comput Ind* 164:104212. <https://doi.org/10.1016/j.compind.2024.104212>
- Gieß A, Hupperz M, Schoormann T, Möller F (2024). What Does it Take to Connect? Unveiling Characteristics of Data Space Connectors. In: proceedings of the 57th Hawaii International Conference on System Sciences, Hawaii, USA.
- Gray J (1996) Evolution of data management. *Computer* 29(10):38–46. <https://doi.org/10.1109/2.539719>
- Guo J, Cheng Y, Wang D, Tao F, Pickl S (2021) Industrial Dataspace for smart manufacturing: connotation, key technologies, and framework. *Int J Prod Res*. <https://doi.org/10.1080/00207543.2021.1955996>
- Harby AA, & Zulkernine F. (2022) From Data Warehouse to Lakehouse: A Comparative Review. In 2022 IEEE international conference on big data (Big Data) (pp. 389–395). IEEE. <https://doi.org/10.1109/BigData55660.2022.10020719>

- Hechler E, Weihrauch M, Wu Y (2023b) Data fabric and data mesh in a hybrid cloud landscape. In: Hechler E, Weihrauch M, Wu Y (eds) *Data fabric and data mesh approaches with AI*. Apress, New York. https://doi.org/10.1007/978-1-4842-9253-2_12
- Hechler E, Weihrauch M, Wu Y (2023c) Data Fabric and Data Mesh Use Case Scenarios. In: Hechler E, Weihrauch M, Wu Y (eds) *Data Fabric and Data Mesh Approaches with AI*. Apress, New York
- Hechler E, Weihrauch M, & Wu Y. (Eds.) (2023a) *Data fabric and data mesh approaches with AI*. Apress. <https://doi.org/10.1007/978-1-4842-9253-2>
- Hechler E, Weihrauch M, & Wu Y. (2023d) Evolution of data architecture. In E. Hechler, M. Weihrauch, & Y. Wu (Eds.), *Data Fabric and Data Mesh Approaches with AI* (pp. 3–15). Apress. https://doi.org/10.1007/978-1-4842-9253-2_1
- Hechler E, Weihrauch M, & Wu Y. (2023e) Key Data fabric and data mesh capabilities. In: E. Hechler, M. Weihrauch, Catherine (Eds) *Data Fabric and Data Mesh Approaches with AI*. Apress. New York
- Hechler E, Weihrauch M, & Wu Y. (2023f) Terminology: data fabric and data mesh. In *Data Fabric and Data Mesh Approaches with AI: A Guide to AI-based Data Cataloging, Governance, Integration, Orchestration, and Consumption*. Springer. Berlin
- Hutterer A, & Krumay B. (2022) Integrating heterogeneous data in dataspace - a systematic mapping study. In *PACIS 2022 proceedings*. <https://aisel.aisnet.org/pacis2022/222>
- Hutterer A, Krumay B, Mühlburger M (2023) What constitutes a data space? conceptual clarity beyond technical aspects. In *29th Americas conference on information systems (AMCIS)*. https://aisel.aisnet.org/amcis2023/eco_systems/eco_systems/5
- IBM Data Fabric. (2025). *Data-fabric-lösungen*. <https://www.ibm.com/de-de/data-fabric>
- Jarke M, Otto B, Ram S (2019) Data sovereignty and data space ecosystems. *Bus Inf Syst Eng* 61(5):549–550. <https://doi.org/10.1007/s12599-019-00614-2>
- Joshi D, Pratik S, & Rao MP (2021) Data governance in data mesh infrastructures: the Saxo Bank case study
- Kashyap, R. (2021). Silicon lifecycle management (SLM) with in-chip monitoring. In *2021 IEEE International Reliability Physics Symposium (IRPS)* (pp. 1–4). IEEE. <https://doi.org/10.1109/IRPS46558.2021.9405187>
- Kotsev A, Minghini M, Tomas R, Cetl V, Lutz M (2020) From spatial data infrastructures to data spaces—a technological perspective on the evolution of European SDIs. *ISPRS Int J Geo Inf* 9(3):176
- Krystek M, Mazurek C, Morzy M, & Pukacki J (2023) Introducing Data Mesh Paradigm for Smart City Platforms Design. In: proceedings of the 56th Hawaii International Conference on System Sciences, Hawaii, USA.
- Kuftinova NG, Ostroukh AV, Filippova NA, Gaevskii VV, Podgorny AV (2022b) Integration of scalable IT architectures on the basis of data fabric technology. *Russ Eng Res* 42(11):1199–1202. <https://doi.org/10.3103/S1068798X22110144>
- Kuftinova NG, Maksimychev OI, Ostroukh AV, Volosova AV, & Matukhina EN (2022) Data fabric as an effective method of data management in traffic and road systems. In *2022 systems of signals generating and processing in the field of on board communications* (pp. 1–4). IEEE. <https://doi.org/10.1109/IEEECONF53456.2022.9744402>
- Kuftinova NG, Ostroukh AV, Maksimychev OI, Pronin CB, & Yadav AK (2023) Implementation of the data fabric architecture as a sustainable development of industrial platform technologies in road transport systems. In *2023 systems of signals generating and processing in the field of on board communications*. symposium conducted at the meeting of IEEE
- Kuhrmann M, Fernández DM, Daneva M (2017) On the pragmatic design of literature studies in software engineering: An experience-based guideline. *Empir Software Eng* 22(6):2852–2891. <https://doi.org/10.1007/s10664-016-9492-y>
- Kumari S, Muthulakshmi P (2022) Transformative effects of big data on advanced data analytics: open issues and critical challenges. *J Comput Sci* 18(6):463–479. <https://doi.org/10.3844/jcssp.2022.463.479>
- Legner C, Pentek T, Otto B (2020) Accumulating design knowledge with reference models: insights from 12 years’ research into data management. *J AIS* 21(3):735–770. <https://doi.org/10.17705/1jais.00618>
- Li X, Yang M, Xia X, Zhang K, & Liu K. (2022) A distributed data fabric architecture based on metadata knowledge graph. In *2022 5th international conference on data science and information technology (DSIT)* (pp. 1–7). IEEE. <https://doi.org/10.1109/DSIT55514.2022.9943831>

- Lim KL, Whitehead J, Jia D, Zheng Z (2021) State of data platforms for connected vehicles and infrastructures. *Commun Transp Res* 1:100013. <https://doi.org/10.1016/j.commtr.2021.100013>
- Liu K, Yang, M, Li X, Zhang K, Xia X, Yan H (2022) M-Data-Fabric: A Data Fabric System Based on Metadata. In *2022 IEEE 5th international conference on big data and artificial intelligence (BDAI)* (pp. 57–62). IEEE. <https://doi.org/10.1109/BDAI56143.2022.9862807>
- Loukiala A, Joutsenlahti J-P, Raatikainen M, Mikkonen T, Lehtonen T (2021) Migrating from a Centralized Data Warehouse to a Decentralized Data Platform Architecture. In: Ardito L, Jedlitschka A, Morisio M, Torchiano M (eds) *Product-Focused Software Process Improvement*. Springer International Publishing, Berlin. https://doi.org/10.1007/978-3-030-91452-3_3
- Machado I, Costa C, Santos MY (2022a) Advancing data architectures with data mesh implementations. In: de Weerd J, Polyvyanyy A (eds) *Lecture notes in business information processing Intelligent Information Systems*. Springer International Publishing., Berlin
- Machado I, Costa C, Santos MY (2022b) Data mesh: concepts and principles of a paradigm shift in data architectures. *Procedia Computer Sci* 196:263–271. <https://doi.org/10.1016/j.procs.2021.12.013>
- Machado I, Costa C, Santos MY (2021) Data-driven information systems: the data mesh paradigm shift. *Proceedings of the 29th International Conference on Information Systems Development*, Valencia, Spain.
- Mazumdar S, Seybold D, Kritikos K, Verginadis Y (2019) A survey on data storage and placement methodologies for Cloud-Big Data ecosystem. *J Big Data* 6(1):15. <https://doi.org/10.1186/s40537-019-0178-3>
- Möller F, Jussen I, Springer V, Gieß A, Schweihoff JC, Gelhaar J, Guggenberger T, Otto B (2024) Industrial data ecosystems and data spaces. *Electron Mark* 34(1):41. <https://doi.org/10.1007/s12525-024-00724-0>
- Nambiar A, Mundra D (2022) An overview of data warehouse and data lake in modern enterprise data management. *Big Data Cognitive Computing* 6(4):132. <https://doi.org/10.3390/bdcc6040132>
- Nanthaamornphong A, Holmes J, Asawateera P (2020) A Case Study: Phuket City Data Platform. In *2020 17th International conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON)* (pp. 717–722). IEEE. <https://doi.org/10.1109/ECTI-CON49241.2020.9158101>
- Nast M, Rother B, Golatowski F, Timmermann D, Leveling J, Olms C, & Nissen C (2020) Work-in-progress: towards an international data spaces connector for the internet of things. In *2020 16th IEEE International conference on factory communication systems (WFCS)* (pp. 1–4). IEEE <https://doi.org/10.1109/WFCS47810.2020.9114503>
- Nema W (2010) An Information Services Architecture Model Based on Data Processing Stages. *IT Prof* 12(1):46–50. <https://doi.org/10.1109/MITP.2010.22>
- Netflix Data Mesh. (2022). *Data Mesh — A Data Movement and Processing Platform @ Netflix*. <https://netflixtechblog.com/data-mesh-a-data-movement-and-processing-platform-netflix-1288bcab2873>
- Opriel S, Möller F, Strobel G, Otto B (2024) Data Sovereignty in Inter-organizational Information Systems. *Advance online publication, Bus Inf Syst Eng*. <https://doi.org/10.1007/s12599-024-00893-4>
- Östberg P-O, Vyhmeister E, Castañé GG, Meyers B, van Noten J (2022) Domain models and data modeling as drivers for data management: The ASSISTANT data fabric approach. *IFAC-PapersOnLine* 55(10):19–24. <https://doi.org/10.1016/j.ifacol.2022.09.362>
- Otto B, Bärenfänger R, Steinbuß S (2015). Digital business engineering: Methodological foundations and first experiences from the field. In *BLLED 2015 Proceedings. 13*. <https://aisel.aisnet.org/bled2015/13>
- Paik H-Y, Xu X, Bandara HMND, Lee SU, Lo SK (2019) Analysis of data management in blockchain-based systems: from architecture to governance. *IEEE Access* 7:186091–186107. <https://doi.org/10.1109/ACCESS.2019.2961404>
- Pampus J, Jahnke B-F, Quensel R (2022) Evolving data space technologies: lessons learned from an ids connector reference implementation. In: Margaria T, Steffen B (eds) *Leveraging applications of formal methods, verification and validation*. Practice, Springer Nature. Switzerland. https://doi.org/10.1007/978-3-031-19762-8_27
- Papageorgopoulos N, Vergeti D, Politi E, Ntalaperas D, Papageorgiou XS, Karvounis M, Stoitsis G, Rakic B, Kotlar M, Petrov S, & Kunchev A. (2022). An agri-food data platform for food safety and certification. In *2022 international conference on computational science and computational intelligence (CSCI)* (pp. 634–639). IEEE. <https://doi.org/10.1109/CSCI58124.2022.00117>
- Pongpech WA (2023) A distributed data mesh paradigm for an event-based smart communities monitoring product. *Procedia Computer Sci* 220:584–591. <https://doi.org/10.1016/j.procs.2023.03.074>

- Pörtner L, Möske R, Riel A (2023) Data ecosystem for industrial product- service systems (IPS2) based on a decentralized data architecture. *Procedia CIRP* 119:1228–1233. <https://doi.org/10.1016/j.procir.2023.02.190>
- Amazon Redshift. (2025). <https://aws.amazon.com/de/redshift/>
- Reuver M de, Ofé H, Agahari W, Abbas AE, Zuiderwijk A (2022) The openness of data platforms: A research agenda. In proceedings of the 1st international workshop on data economy (pp. 34–41). ACM. <https://doi.org/10.1145/3565011.3569056>
- Rieyan SA, News MRK, Rahman AM, Khan SA, Zaarif STJ, Alam MGR, Hassan MM, Ianni M, Fortino G (2024) An advanced data fabric architecture leveraging homomorphic encryption and federated learning. *Inf Fusion* 102:102004. <https://doi.org/10.1016/j.inffus.2023.102004>
- Roach CJ (1962) Symposium on clinical drug evaluation and human pharmacology. XVIII. *Data Manag Clin Pharmacol Ther* 3:682–684. <https://doi.org/10.1002/cpt196235682>
- Saini A, Sharma V, Kumar M, Dey A, & Tripathy D (2023) Security, privacy, and authentication framework for web-driven data fabric. In: V. Sharma, B. Balusamy, J. Thomas, L. Atlas (Eds.), *Data Fabric Architectures*. De Gruyter. pp. 183–198 <https://doi.org/10.1515/9783111000886-010>
- Scheffler M, Aeschlimann M, Albrecht M, Bereau T, Bungartz H-J, Felser C, Greiner M, Groß A, Koch CT, Kremer K, Nagel WE, Scheidgen M, Wöll C, Draxl C (2022) Fair data enabling new horizons for materials research. *Nature* 604(7907):635–642. <https://doi.org/10.1038/s41586-022-04501-x>
- SCSN. (2023). *Smart Connected Supplier Network*. <https://smart-connected.nl/en>
- Sedek KA, Osman MN, Omar MA, Wahab MHA, Idrus SZS (2021) Smart agro E-marketplace architectural model based on cloud data platform. *J Phys: Conf Ser* 1874(1):12022. <https://doi.org/10.1088/1742-6596/1874/1/012022>
- Šestak M, Copot D (2023) Towards trusted data sharing and exchange in agro-food supply chains: design principles for agricultural data spaces. *Sustainability* 15(18):13746. <https://doi.org/10.3390/su151813746>
- Sharma K, Shankar A, Prabhishek Singh A, Shankar PS (2021) Information security assessment in big data environment using fuzzy logic. *J Cybersecur Inf Manag* 5(1):29–42. <https://doi.org/10.54216/JCIM.050103>
- Sherman R (2015) Data architecture In *Business Intelligence Guidebook*. Elsevier. Amsterdam pp 107–142 <https://doi.org/10.1016/B978-0-12-411461-6.00006-X>
- Snowflake (2025). *Snowflake - A platform like no other*. <https://www.snowflake.com/en/>
- Stonebraker M, Çetintemel U (2018) One size fits all": an idea whose time has come and gone. In: Massachusetts Institute of Technology & M. L. Brodie (Eds) *Making Databases Work: the Pragmatic Wisdom of Michael Stonebraker*. Association for Computing Machinery. <https://doi.org/10.1145/3226595.3226636>
- Strengholt P (2023) *Data Management at Scale: Modern data architecture with data mesh and data fabric* (Second edition). O'Reilly
- Tardieu H (2022) Role of Gaia-X in the European data space ecosystem. In *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage*. Springer International Publishing. Cham.
- Theodorou V, Gerostathopoulos I, Alshabani I, Abelló A, Breitgand D (2021) Medal: an AI-driven data fabric concept for elastic cloud-to-edge intelligence. In: Barolli L, Woungang I, Enokido T (eds) *Advanced information networking and applications*. Springer International Publishing, Berlin Heidelberg
- Tian Y (2022) Ai-assisted dynamic modeling for data management in a distributed system. *J Inter Net* 22(Supp05):2147002. <https://doi.org/10.1142/S0219265921470022>
- Truong NB, Sun K, Lee GM, Guo Y (2020) GDPR-compliant personal data management: a blockchain-based solution. *IEEE Trans Inform Forensic Secur* 15:1746–1761. <https://doi.org/10.1109/TIFS.2019.2948287>
- Turland M, Slade P (2020) Farmers' willingness to participate in a big data platform. *Agribusiness* 36(1):20–36. <https://doi.org/10.1002/agr.21627>
- vom Brocke J, Simons A, Riemer K, Niehaves B, Plattfaut R, Cleven A (2015) Standing on the shoulders of giants: challenges and recommendations of literature search in information systems research. *CAIS* 37(1):9. <https://doi.org/10.17705/1CAIS.03709>
- von Scherenberg F, Hellmeier M, Otto B (2024) Data sovereignty in information systems. *Electron Mark* 34(1):15. <https://doi.org/10.1007/s12525-024-00693-4>
- Wang RY, Strong DM (1996) Beyond accuracy: what data quality means to data consumers. *J Manag Inf Syst* 12(4):5–33. <https://doi.org/10.1080/07421222.1996.11518099>

- Webster J, & Watson RT (2002) Analyzing the past to prepare for the future: writing a literature review. *MIS Quarterly*, xiii–xxiii
- Wider A, Verma S, Akhtar A (2023) Decentralized data governance as part of a data mesh platform: concepts and Approaches. In *2023 IEEE international conference on web services (ICWS)* (pp. 746–754). IEEE. <https://doi.org/10.1109/ICWS60048.2023.00101>
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, Da Silva Santos LB, Bourne PE et al (2016) The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3(1):1–9
- Wrembel R (2023) Data integration revitalized: from data warehouse through data lake to data mesh. In: Strauss C, Amagasa T, Kotsis G, Tjoa AM, Khalil I (eds) *Lecture notes in computer science database and expert systems applications*. Springer Nature, Switzerland. https://doi.org/10.1007/978-3-031-39847-6_1
- Yao Z, Zhong Y, Liao Q, Wu J, Liu H, Yang F (2021) Understanding human activity and urban mobility patterns from massive cellphone data: platform design and applications. *IEEE Intell Transp Syst Mag* 13(3):206–219. <https://doi.org/10.1109/MITS.2019.2962146>
- Zhao M, Kumar A, Ali GMN, & Chong PHJ (2016) A cloud-based network architecture for big data services. In *2016 IEEE 14th intl conf on dependable, autonomic and secure computing, 14th intl conf on pervasive intelligence and computing, 2nd intl conf on big data intelligence and computing and cyber science and technology congress(DASC/PiCom/DataCom/CyberSciTech)*. Symposium conducted at the meeting of IEEE
- Zhu JY, Tang B, Li VO (2019) A five-layer architecture for big data processing and analytics. *Int J Big Data Intell* 6(1):38. <https://doi.org/10.1504/IJBDI.2019.097399>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.