

Systematic Literature Review on Data Mesh Architecture: A Bibliometric Analysis

Seyed Hassan Mousavi 

Department of Management, Na.C.,
Islamic Azad University,
Najafabad, Iran.
s.mousavi2133@iau.ac.ir

Akbar Nabiollahi* 

Faculty of Computer Engineering,
Na.C., Islamic Azad University,
Najafabad, Iran.
AK.Nabioollahi@iau.ac.ir

Rasoul Asgarpour 

Department of Management, Na.C.,
Islamic Azad University,
Najafabad, Iran.
Rasoulasgarpour@iau.ac.ir

Shervan Fekri-Ershad 

Faculty of Computer Engineering, Na.C.,
Islamic Azad University,
Najafabad, Iran.
fekriershad@iau.ac.ir

Abbas Zamani 

Department of Management, Na.C.,
Islamic Azad University,
Najafabad, Iran.
abbas.zamani61@iau.ac.ir

Received: 14 January 2025 – Revised: 12 July 2025 - Accepted: 14 July 2025

Abstract—With the rapid growth and increasing complexity of data, organizations struggle to manage it effectively, thereby revealing the limitations of traditional data architectures. Data Mesh offers a solution through a decentralized, domain-oriented approach that treats data as a product and is supported by federated governance models and self-service infrastructure. Despite its increasing adoption, discussions of its governance within academic circles remain scattered and unstructured. The main objectives of this article are to provide an in-depth overview, share valuable insights into research paths and trends within the field of Data Mesh architecture, and synthesize conclusions from previous studies. This is due to limited domestic research, the fragmented nature of existing studies, and the absence of systematic analysis in this field. Through a comprehensive literature review, this study aims to elucidate key Data Mesh concepts, including distributed data architectures and decentralized governance. By highlighting significant publications, authors, and journals influencing the discourse on Data Mesh, it establishes a benchmark for measuring the breadth and depth of research in this area. One of the aims of this study is to enhance future research trajectories addressing relevant issues. Using VOSviewer software, the methodology integrates bibliometric analysis and systematic review to identify and examine research trends in Data Mesh architecture. Purposive sampling was employed to identify relevant research from the Scopus and WOS databases, with Google Scholar incorporated as a supplementary database for the study's statistical population. The analysis of keyword co-occurrence networks constitutes one of the study's main findings. Furthermore, four clusters based on citation counts were identified in the co-citation network of sources, with the Computer Science Information Systems and Computer Science Theory and Methods journals forming the largest clusters. A key finding of this research is the identification of significant indicators of Data Mesh architecture through an extensive literature analysis aimed at developing a framework for decentralized data governance. This study provides practitioners and policymakers with evidence-based recommendations for effectively implementing Data Mesh principles. The study also promotes interdisciplinary collaboration by demonstrating the connections between Data Mesh and related domains such as blockchain. Finding future research avenues on the application of decentralized technology in data governance-related concerns is one of the most important results.

Keywords: Data Mesh, Decentralized Data Governance, Distributed Data Architecture, Data Management.

Article type: Research Article



© The Author(s).

Publisher: ICT Research Institute

* Corresponding Author

I. INTRODUCTION

The unprecedented growth in data volume, variety, and velocity has revealed major limitations in traditional centralized data architectures, particularly concerning scalability, flexibility, and operational efficiency. Conventional data governance models, characterized by centralized authority over policies and standards, often hinder organizational agility and responsiveness. Data governance is defined as a comprehensive framework of methodologies, policies, and processes that organizations employ to ensure the proper management, protection, and utilization of their data assets, thereby aligning data-centric activities with organizational goals to support effective decision-making, regulatory compliance, and risk mitigation. Among the most critical issues associated with data governance—particularly within information technology organizations—are those related to data ownership, accountability, responsibility, and challenges concerning control mechanisms [2, 32, 33].

In response to these limitations and challenges, the Data Mesh architecture—introduced by Dehghani in 2019 [6, 7]—is considered an innovative decentralized approach to large-scale data governance, applying domain-driven design and self-serve data infrastructure. This architecture focuses on transitioning from centralized data warehouses and data lakes to a model that delegates data ownership to organizational domains, enabling individual teams to manage, provide, and consume data as a product. It emphasizes four core principles—domain-oriented data ownership, data as a product, self-serve data infrastructure, and federated computational governance—providing a flexible and scalable framework for modern data needs [6, 7]. This approach ensures that policies concerning data ownership, access control, compliance, and lifecycle management are managed close to where the data resides. Although academic engagement with Data Mesh remains in its early stages, it is steadily increasing. Increasingly, studies are examining its conceptual foundations, governance challenges, technical limitations, and potential integrations with technologies like blockchain and metadata catalogs.

Reviewing and analyzing the literature provides a structured method for synthesizing existing knowledge and identifying key trends, research gaps, and future research directions. Additionally, scientometric analysis is a quantitative approach used to examine existing literature by evaluating publication patterns, citation networks, and the impact of key papers and authors in this field. Through scientometric methods, researchers can assess the intellectual structure of this research area, track its development over time, and identify influential contributions [35]. This paper aims to identify and analyze trends in studies related to Data Mesh architecture through a combination of systematic literature review and bibliometric methods, employing the VOSviewer software to provide a comprehensive understanding of the current state of research. This analysis offers valuable insights into emerging research trends, identifies influential papers, gaps in the literature, and potential directions for future studies, acting as a resource for researchers, practitioners, and organizations looking to explore and adopt the

principles of this architecture in their data strategies [35].

Through this literature review and bibliometric analysis, the study identifies key aspects such as understanding and integrating the emerging paradigm in data governance, recognizing research gaps, and synthesizing scattered knowledge on Data Mesh architecture. It also clarifies the alignment of data ownership with accountability, enhances operational efficiency and scalability, and addresses political and ethical considerations in implementing decentralized systems [2, 4, 35]. Bibliometric analysis helps identify under-researched areas or technologies with potential for further exploration [35]. This study can guide future research on applying blockchain technology to develop decentralized data governance frameworks or using machine learning to automate data governance tasks [2, 4, 33].

In this research, using literature review and bibliometric analysis, we aim to find the key indicators of data mesh architecture to respond to each of the challenges mentioned in the development of a decentralized data governance framework. This architecture, by distributing data governance tasks and roles with innovative solutions, allows individuals and users to have more control over their data. Each unit is responsible for managing and administering data in its domain, following its own set of policies and guidelines, ensuring consistency and collaboration between units through a central section [2, 8].

Reference [4] propose a conceptual governance framework within a Data Mesh platform, focusing on the automation of policies via access controls, though issues of accountability and ownership remain inadequately explored. Similarly, [17] through interviews with industry experts, identify key motivational factors and transition challenges, particularly in relation to change management and domain readiness, as hurdles in shifting from centralized to federated governance models.

Additional contributions focus on aspects like privacy and data quality. Reference [15] investigate embedding privacy engineering principles within federated governance frameworks, while [37] identify 47 subcomponents of data quality management relevant to decentralized architectures. Hendriks [42], through a systematic review of gray literature, reveals notable gaps in standardization, governance of the data product lifecycle, and the maturity of associated tools.

Blockchain-based implementations have emerged as promising extensions for decentralization. Reference [13] advocate for a blockchain-enhanced governance model using smart contracts to automate data access validation and policy enforcement. Also [9] incorporate non-fungible tokens (NFTs) to secure time-bound data ownership, enhancing both transparency and auditability. Voß [26] takes this further by transforming datasets into digital assets, enabling cross-domain reuse through programmable contracts.

Unlike previous studies that primarily concentrated on architectural elements or conceptual governance, this paper employs a combination of qualitative thematic coding and bibliometric analysis to identify

governance-specific indicators from a wider dataset. This integrated approach supports the creation of a decentralized data governance framework, rooted in validated principles and measurable criteria, providing a clear and practical roadmap for implementation.

Despite these advancements, much of the existing research is conceptual or exploratory. Few studies provide comprehensive frameworks aligning Data Mesh principles with established governance models like the DGI framework. Addressing this gap, the present study synthesizes themes around accountability, ownership, and control mechanisms, aligning them with governance best practices and actionable guidelines.

The research questions guiding this review include: 1- What are the key topics and trends in data mesh literature? 2- Who are the influential authors and institutions in this field? 3- What gaps exist in current research and where are future opportunities? 4- What are the challenges and limitations for implementing a data mesh? 5- What are the key and influential indicators of data mesh architecture for implementing to apply for developing a decentralized data governance framework?

The rest of this paper is structured as follows: Section 2 introduces literature review of related works, Section 3 explains the proposed research methodology and related basic concepts, Section 4 describes the analysis of findings, Section 5 reports conclusions and some parts are dedicated to the discussion.

II. LITERATURE REVIEW

The Data Mesh architecture, introduced by Dehghani in 2019 [6, 7], represents an innovative decentralized approach to large-scale data governance through the application of data product thinking and domain-oriented decentralization of data ownership. Within this architecture, organizational data assets are treated as products and made available to data consumers for efficient and independent use. Domain teams—being closest to the data sources—create, maintain, and manage these data products. As domains consume one another's products, a networked mesh of interconnected data products emerges. To ensure interoperability and compliance with both internal and external policies and regulations, a federated governance team defines and enforces shared standards and compliance requirements across the data mesh. Finally, a self-service data platform supports each domain by providing the underlying infrastructure and tools necessary to build, sustain, and manage their data products effectively [2, 6, 7].

Several recent studies have investigated research questions related to Data Mesh architecture. Reference [4] presents a conceptual model outlining the key concepts of this architecture and discusses various approaches to steering governance through platform-based tools. However, that study focuses primarily on data access control and does not address issues of data ownership and accountability. Reference [14] explores motivational factors for adopting Data Mesh,

associated challenges, implementation strategies, business impacts, and potential design patterns. To bridge this research gap, the authors employed a qualitative research approach based on 15 semi-structured interviews with industry experts, revealing that organizations encounter difficulties when transitioning to the federated governance model inherent in the Data Mesh paradigm. Similarly, Reference [13] addresses several implementation challenges and solutions, including metadata management, technological support, and blockchain integration. Metadata management within federated governance encompasses the administration of data schemas, semantic models, and data lineage.

Reference [37], using a meta-synthesis method and 268 sources, identified 12 main categories and 47 subcategories related to data quality management within data governance through content analysis. Reference [41] identified the essential requirements for establishing big data governance in organizations, recognizing 48 fundamental themes derived from text analysis and interviews. The ability of these organizations to excel in informed decision-making, strategic planning, analytics, effective governance, and administrative transparency depends on the vast amounts of data gathered from both internal and external sources, which must be processed and safeguarded through robust big data governance practices. After incorporating expert feedback and revisions, the final count included 36 themes—comprising 24 antecedents and 12 consequences. To validate the identified themes, the fuzzy Delphi method was applied in two stages, and by reviewing the defuzzified mean values from both stages, both the antecedents and consequences were confirmed.

Reference [13] discusses a blockchain-based implementation of a Data Mesh platform, serving as a practical validation of the theoretical proposal for federated governance within the Data Mesh architecture. This implementation aims to support decentralized data domains and strengthen the governance framework required for the successful adoption of the Data Mesh approach. Reference [15] explores the integration of blockchain technology within the Data Mesh framework to address challenges of data ownership and governance. It introduces a novel mechanism that employs Non-Fungible Tokens (NFTs) to facilitate secure and transparent data ownership transfers. This mechanism enables the minting of time-based NFTs, ensuring that stakeholders can access specific portions of the Data Mesh as needed. The combination of these technologies aims to establish a user-centric, secure, and efficient data management system for decentralized architectures. Similarly, Reference [30] proposes a distributed blockchain-based framework inspired by the Data Mesh architecture to address challenges such as poor scalability, ambiguous governance, isolated data silos, and duplication of effort. The proposed framework leverages blockchain's decentralized nature to enable efficient

and transparent dataset sharing across enterprise business domains.

Reference [11] discusses the integration of blockchain technology within a Data Mesh architecture, specifically through the implementation of a blockchain-powered metadata catalog. This catalog enhances data management by enabling federated control, ensuring immutability, and promoting transparency across distributed data domains. Reference [68] identifies key challenges

associated with applying the Data Mesh approach to the development of a data platform tailored for smart cities. It also provides guidelines for the successful introduction of the Data Mesh at sociological, technological, and infrastructural levels. Moreover, the study highlights the applicability of the Data Mesh paradigm in the context of digital twins—a promising vision for the future of smart city services.

Some key studies on data mesh architecture are summarized in the table 1.

TABLE I. REVIEW OF KEY RESEARCHES ON DATA MESH ARCHITECTURE

Research Title	Research Gap	Research Findings	Relevance to Current Research
Implementation of Federated Governance in Data Mesh Architecture[13]	Lack of technology support for data mesh adoption. Need for formal governance definition in data mesh.	An implementation based on blockchain technology to validate the idea in practice.	Critique of existing implementations and emphasis on the need for technological support and formal governance
Security and Ownership in Data Mesh[9]	Lack of technology support for data mesh adoption. Need for formal governance definition in data mesh	An innovative algorithm that leverages Blockchain and Non-Fungible Tokens (NFTs) to facilitate secure and transparent data ownership transfers.	Critiques existing frameworks by proposing a new algorithmic approach using blockchain and NFTs.
Data Mesh Architecture: From Theory to Practice[67]	This article fails to address the existing research gap.	The key components of data mesh architecture. Discusses design decisions for implementing data mesh.	Emphasizes on domain ownership and independent data products.
Privacy Engineering in Data Mesh: Towards a Framework for Decentralized Data Privacy[15]	Intersection of privacy engineering and data mesh dimensions Guidelines for stakeholders to embed better data privacy controls	Intersection of privacy engineering and data mesh dimensions. Guidelines for stakeholders to embed better data privacy controls.	Critiques existing frameworks for not integrating privacy principles. Proposes a comprehensive approach to maintaining privacy across the data mesh architecture.
Decentralized Data Governance as Part of a Data Mesh Platform: Concepts and Perspectives[4]	Lack of coherence in data mesh concept interpretations. Minimal work on governance facilitation in data mesh platforms.	Policies based on access control mechanisms. A starting point for the maturation of technology needed for decentralized data governance.	The presented concepts and approaches help design a data mesh platform, serving as a guide on how to interpret data mesh concepts.
Data Mesh: A Grey Literature Review[2]	Scarcity of academic research on data mesh. Need for standardization of data products and interoperability.	Research challenges such as standardization, methods and tools for architecture development and operation, data product lifecycle.	Systematically reviews 114 articles to identify key design principles, architectural components, organizational roles, benefits, and concerns.
Data Mesh: The Best Way to Avoid Data Mes[7]	This article did not make any reference to the research gap.	Motivational indicators for introducing data mesh and the corresponding challenges.	Providing the best solutions and business impacts - with a qualitative view of motivational indicators, challenges, best solutions.
Data Mesh: Motivational Factors, Challenges, and Best Practices[17]	Lack of quantitative evidence for findings and archetypes. Need for deeper exploration in small organizations.	Identified challenges in transitioning to data mesh governance. Derived best practices for successful data mesh adoption.	Emphasizing on domain-oriented responsibilities and treating data as a product for improved accessibility and quality.
Advancing Data Architectures with Data Mesh Implementations[5]	Lack of scientific foundation	Assesses the concept using a case study that demonstrates the value and advantages of this kind of data architecture.	suggests a technical framework for implementing a data mesh
Identifying Alternatives and Deciding Factors for a Data Mesh Architecture[26]	Lack of Factors to consider before implementing a data mesh	five key factors to consider before implementing a data mesh.	Finding Options and Choosing Elements for a Data Mesh System

In previous studies, the focus has been on the challenges of implementing Data Mesh architecture,

the necessities and motivations for its application in organizations, and the discussion of its advantages,

disadvantages, related challenges, and architectural components [2, 4]. The present study aims to identify the key topics and trends in the Data Mesh architecture literature, determine the most influential authors and institutions in this field, and highlight existing research gaps to pinpoint future research opportunities. One of the primary objectives of this article is to extract the key and influential indicators of Data Mesh architecture through a literature review and bibliometric analysis, which will serve as the foundation for developing a decentralized data governance framework.

III. RESEARCH METHODOLOGY

By integrating findings from previous studies to advance a specific research domain, this study employs a combination of a systematic literature review and bibliometric analysis methods[57]. Bibliometric techniques, characterized by their analytical precision, can be used as a complementary approach to qualitative literature reviews. Bibliometric analysis enables a structured and quantitative examination of a large body of research, providing an overarching view of existing scholarship. It captures key indicators such as the annual distribution of publications, the most cited sources and authors, cumulative growth trends, keyword networks, co-occurrence patterns, and co-citation relationships among sources [35].

The statistical population of this study comprises scientific works related to Data Mesh architecture and decentralized data architectures, identified using the keywords listed in Table 2. The selected publications span the years 2010 to 2024 and were retrieved from the electronic databases *Web of Science*, *Scopus*, and *Google Scholar*, with full-text availability online.

To ensure comprehensive coverage, the decision was made to use *Web of Science* and *Scopus* as the two primary scientific databases, while incorporating *Google Scholar* as a supplementary source. The inclusion of *Google Scholar* allowed the integration of

non-peer-reviewed and non-formally published materials to enrich and complement the findings derived from the main scientific databases.

Following the frameworks proposed in [45] and [56], the process of selecting and reviewing articles consists of four main stages. The first stage involves constructing the databases according to predefined criteria. The second stage includes identifying 93 published articles based on keyword searches in the *Web of Science* database and 91 articles in the *Scopus* database. A search on *Google Scholar* for the term "Data Mesh" in article titles returned 231 results, demonstrating its value as a supplementary source for the review. Finally, a specific cutoff point will be established to determine the number of articles included in this study.

Based on the titles, keywords, and abstracts, 70 relevant studies were initially identified. After assessing the accessibility of the articles and reviewing their full texts, various sections of each study—including the title, abstract, content, methodology, and references—were evaluated. Ultimately, 60 research articles were selected for inclusion. In the third stage, bibliometric and systematic review methods were applied to analyze the selected articles. In the final stage, the results were reported, and recommendations were formulated.

To ensure the inclusion of high-quality and relevant studies, the inclusion criteria consist of peer-reviewed articles, conference papers, and reviews focused on data mesh architecture or related concepts, published between 2019 (the year data mesh architecture was introduced) and 2024. The exclusion criteria include articles not written in English, articles that only briefly mention data mesh without in-depth discussion, and non-scientific sources [35].

Figure 1 illustrates the screening and analysis process of articles in the relevant databases.

TABLE II. KEYWORDS AND SEARCH TERMS

Criteria	Database	Detailed Criteria	Inclusion Criteria	Number of Articles
Keywords and Subjects	WOS	"data mesh" (Titles) OR "decentralized data governance" (Topics) AND "federated computational governance" (Topics) OR "decentralized data architecture" (Topics) OR "decentralized data management" (Topics)	1-English articles 2-Peer-reviewed articles 3- Articles with detailed publication metadata 4- Open Access	93
	Scopus	(TITLE-ABS-KEY ("data mesh") OR TITLE-ABS-KEY ("decentralized data architecture") OR TITLE-ABS-KEY ("distributed data architecture"))		91
	Google Scholar	"data mesh"		231
Document type		Paper		
Study type		Conference paper/ Scientific Research paper/ Review paper		
Time duration		2010-2025		

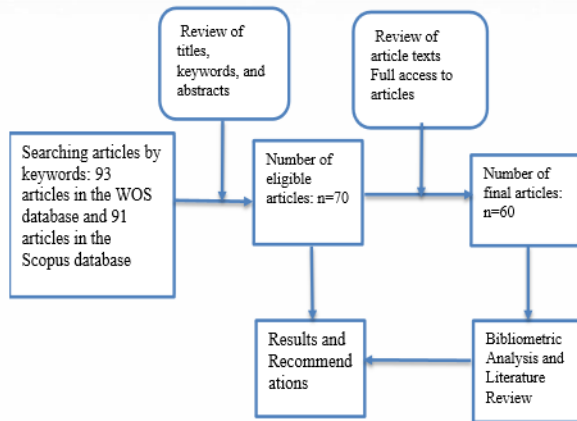


Figure 1. Screening and analysis process of articles in the database

This extensive strategy ensures wider access to relevant literature and encompasses various perspectives on Data Mesh Architecture, Decentralized Data Governance, and Federated Data Management. The selected articles were entered into VOSviewer software for further processing and bibliometric

analysis. Metadata, including authors' names, publication years, titles, abstracts, keywords, and citations, were extracted for analysis. Network visualizations and density maps were created using VOSviewer to gain a better understanding of the connections between various research issues in the field of data mesh and decentralized data architectures. The current status of research in this field is reflected in these graphics. Items in these maps are represented by multicolored dots, with each dot's size reflecting how frequently it appears in the papers. The connected phrase appears more frequently when the dot is larger. Additionally, each dot's color indicates the social or thematic cluster to which it belongs.

The analyses focused on co-authorship networks, identifying collaborations between authors and institutions; co-citation analysis, examining the number of times two articles have been cited together; and keyword analysis, identifying recurring keywords and topics in the literature. Figures 2 and 3 show the co-occurrence analysis, where the unit of analysis is the number of authors' keywords with a minimum of two occurrences.

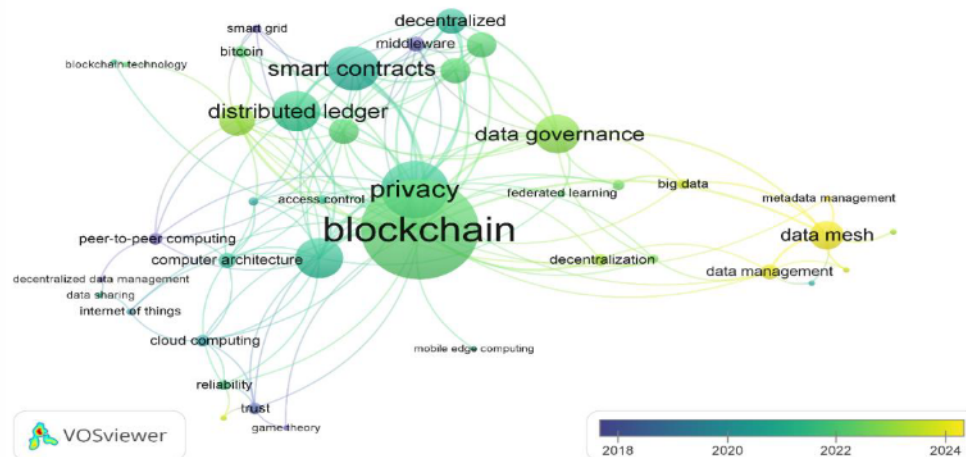


Figure 2. Co-Occurrence Analysis (Unit of Analysis: Number of Authors' Keywords, Minimum number of occurrence keyword=2) in the WOS Database

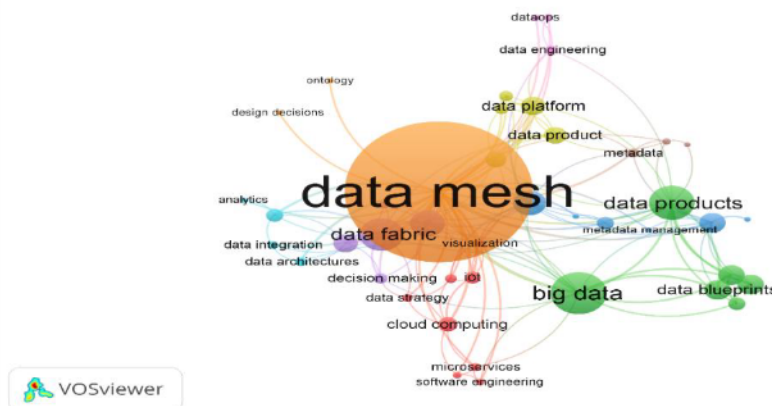


Figure 3. Co-Occurrence Analysis (Unit of Analysis: Number of Authors' Keywords in the Scopus Database)

This study identifies several key authors who have made significant contributions to the development and conceptualization of data mesh architecture. Among these individuals, Zhamak Dehghani is recognized as the founder of the data mesh architecture paradigm. Leading academic institutions and companies, particularly in the USA, Europe, and Australia, are at the forefront of this research. Keyword analysis reveals several recurring themes within the body of literature. Key terms such as decentralized data governance, data ownership, domain-oriented data architecture, and self-serve data platforms dominate this field. These concepts embody the core principles of the architectural paradigm, emphasizing the shift from centralized data governance to a distributed, domain-oriented approach.

IV. ANALYSIS OF FINDING

This section integrates the primary bibliometric and qualitative insights derived from the literature on Data Mesh architecture, relying on the previously outlined screening and coding procedures described in the methodology. The progression of academic interest in data mesh architecture shows a distinct upward trend, particularly from 2019 onward, aligning with Dehghani's introduction of the concept [6, 7]. A bibliometric analysis of Web of Science (WOS) and Scopus databases reveals a compound annual growth rate surpassing 5%, highlighting consistent academic engagement. Between 2010 and 2015, research on decentralized data governance and domain-driven data architecture was scarce and mostly exploratory or conceptual.

From 2016 to 2018, a gradual increase was evident, driven by foundational studies on federated governance and distributed architectures. The turning point came in 2019 when Dehghani's work sparked extensive discussion on domain-oriented data ownership, self-serve infrastructure, and treating data as a product. Between 2020 and 2022, there was a substantial rise in peer-reviewed articles, theses, and industrial case studies, particularly from IEEE, ACM, and Springer platforms. The WOS database listed 93 related publications, Scopus recorded 91 articles, and Google Scholar identified 231 entries, reflecting widespread interest in both academic and gray literature.

From 2022 to 2024, publication rates surged significantly, driven by the incorporation of blockchain, privacy engineering, and smart contracts into the data mesh framework. Influential journals like Information and Organization, Future Internet, and IEEE Access became key contributors, focusing on governance frameworks, interoperability, and automation. This pattern illustrates growing acknowledgment of the shortcomings of centralized systems, alongside increasing collaboration between academia and industry to promote decentralized governance while enhancing scalability, accountability, and compliance.

TABLE III. PERCENTAGE OF WOS CATEGORIES

Web of Science categories	Record count	% of 93
Computer Science Information Systems	39	41.935
Computer Science Theory Methods	33	35.484
Computer Science Artificial Intelligence	16	17.204
Telecommunications	16	17.204
Engineering Electrical Electronic	15	16.129
Computer Science Software Engineering	14	15.054
Computer Science Interdisciplinary Applications	11	11.828
Computer Science Hardware Architecture	9	9.677
Engineering Industrial	3	3.226
Medical Informatics	3	3.226

A. Influential Journals in Data Mesh Architecture

159 articles have been published in 10 journals, with the journal of Computer Sciences Information Systems publishing 39 articles, and the journals of Computer Sciences Theory Methods and Computer Sciences Artificial Intelligence 33 and 16 articles respectively. The tables 3 shows the percentage of WOS categories and percentage of document types in the WOS database. These journals also include the most cited sources respectively.

The bibliometric analysis reveals that while Data Mesh remains an emerging field, certain journals have gained significant influence in shaping its conceptual and technical foundation. According to Web of Science (WOS) and Scopus indexing, the top three contributing journals are 1-Computer Sciences–Information Systems: Accounting for 41.9% of the articles, journals such as *Future Internet* and *Journal of Information Processing & Management* focus on topics like architectural design, federated governance, and domain ownership [4,13]. 2-Computer Sciences–Theory & Methods: Comprising 35.5% of publications, this category delves into conceptual frameworks and methodological innovations within decentralized data systems [2, 5, 29, 68]. 3-Computer Sciences–Artificial Intelligence: Representing 17.2% of contributions, these articles intersect with technologies like blockchain, machine learning, and intelligent metadata systems for automated governance [15, 30, 47].

There is a notable shift toward socio-technical perspectives, merging management science, software engineering, and ethical governance to advance Data Mesh adoption. Journals such as *IEEE Access*, *ACM Computing Surveys*, and *Lecture Notes in Computer Science* are increasingly cited, especially in research exploring real-world applications of Data Mesh through blockchain and smart contract mechanisms [13, 68, 69]. Tabel 4 shows the Percentage of document types in the WOS Database.

TABLE IV. PERCENTAGE OF DOCUMENT TYPES IN THE WOS DATABASE

Document Types	Record Count	of 93 %
Article	50.538	47
Review Article	3	3.226
Book Chapters	1	1.075
Early Access	1	1.075

B. Highly Cited Authors and Publication Trends

The number of highly cited authors is 25. Five authors each account for around 60% of the citations from three research works, and five others (20%) have two research works. The author collaboration index is 2/3, with an average of two authors collaborating on each article. "Casteltrout," "Kumara," "Laurent," "Niyato" and "Dolhopolov" each have 3 articles, and "Santos" "Jeffrey," "Xiong," "Rise," and "Jurdak" each have two articles, making them the most locally cited in these research works. The research priorities of highly cited authors indicate that [4] focuses on examining a conceptual model of key concepts in data mesh architecture and various perspectives for designing a decentralized governance model, primarily focusing on designing a conceptual model for implementing a data mesh platform reference and its application in decentralized data governance. The research priority of [2] is on a systematic review of 114 industrial gray literature articles on data mesh architecture to identify key design principles, architectural components, organizational roles, benefits, and concerns.

The Germany, with over 12 articles, France with 6 articles, England and Ireland with 4 articles, Finland and India, Canada, Brazil and Australia each with 4 articles, have the most articles. The Germany, with 40% of the articles and 20 citations to its articles, ranks first.

Figure 4 shows the percentage of articles in journals in the Scopus database.

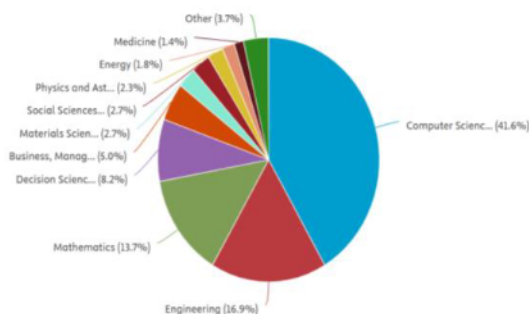


Figure 4. Percentage of articles in journals in the Scopus Database

C. Most Cited Works

Given the novelty of data mesh architecture, there are limited scientific and research sources available. Additionally, since the data mesh concept originates from industry rather than academic environments, resources such as blogs and white papers are widely available [6, 7]. The 5 most cited articles are shown in

table 5. Eight articles used quantitative research methods, and others used qualitative methods. Some articles employed a systematic review approach, while others used design science methods.

In [5], it is emphasized that general models or methods related to the implementation of data mesh architecture are needed. Given the lack of resources for implementing data mesh architecture, this architecture supports the same logic by promoting decentralized domain data ownership as its core, with a team responsible for a specific domain and analytical and operational data within that domain.

In [28], the new paradigm of Data Mesh, a decentralized approach to data architecture that departs from conventional monolithic systems like data lakes, is examined. It demonstrates how Data Mesh uses its fundamental ideas to solve issues with scalability, adaptability, and operational efficiency. In order to facilitate the deployment of Data Mesh in enterprises, the study suggests a conceptual architecture as well as a domain model.

TABLE V. HIGHLY CITED ARTICLES

Article Title	Author	Journal	Thematic Focus	Citation
Data-Driven Information Systems: The Data Mesh Paradigm Shift. Integrated Spatial Databases.	Machado, I.A., Costa, C.A., & Santos, M.Y.(2011)	Integrated Spatial Databases 2021 Computer Science	Domain ownership, self-service platforms	12
Data mesh: a systematic gray literature review	Goedegebuure & Kumara & Driessen. (2024)	ACM Computing Surveys	Research gaps, architectural taxonomy	8
Data Governance in Data Mesh Infrastructure: the Saxo Bank Case Study	Joshi, D., Pratik, S., & Rao, M.(2022)	ICEB 2021 Proceedings (Nanjing, China)	Institutional implementation and policy design	8
Decentralized Data Governance as Part of a Data Mesh Platform: Concepts and Approaches	Wider & Verma & Akhtar.(2023)	IEEE	Federated governance, platform engineering	6
Advancing Data Architectures with Data Mesh Implementations	Machado & Costa & Santos.(2022)	INTELLIGENT INFORMATION SYSTEMS (CAISE FORUM 2022)	Value-driven implementations	5

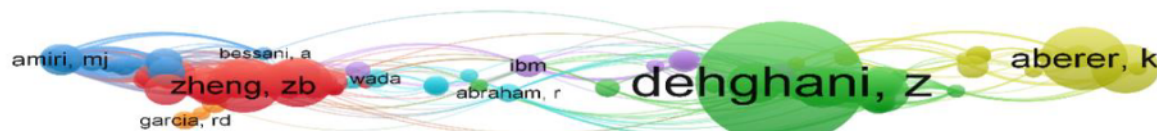


Figure 5. Co-Citation Networks of Authors in the WOS Database

TABLE VI. CLUSTER FORMATION

Cluster number	Thematic Focus	Key Sources	Contribution
1	Core Architecture Foundations	[2, 4, 6, 7]	Introduction of Data Mesh, conceptual models, and federated governance
2	Design & Technical Frameworks	[5, 11, 68, 69]	Blockchain integration, metadata catalogs, technical design practices
3	Governance and Control Mechanisms	[13, 58, 59, 60, 61]	Data ownership, accountability, DGI, and control models
4	Privacy, Regulation, and Interoperability	[15, 30, 64, 61]	Privacy engineering, NFT ownership, decentralized interoperability

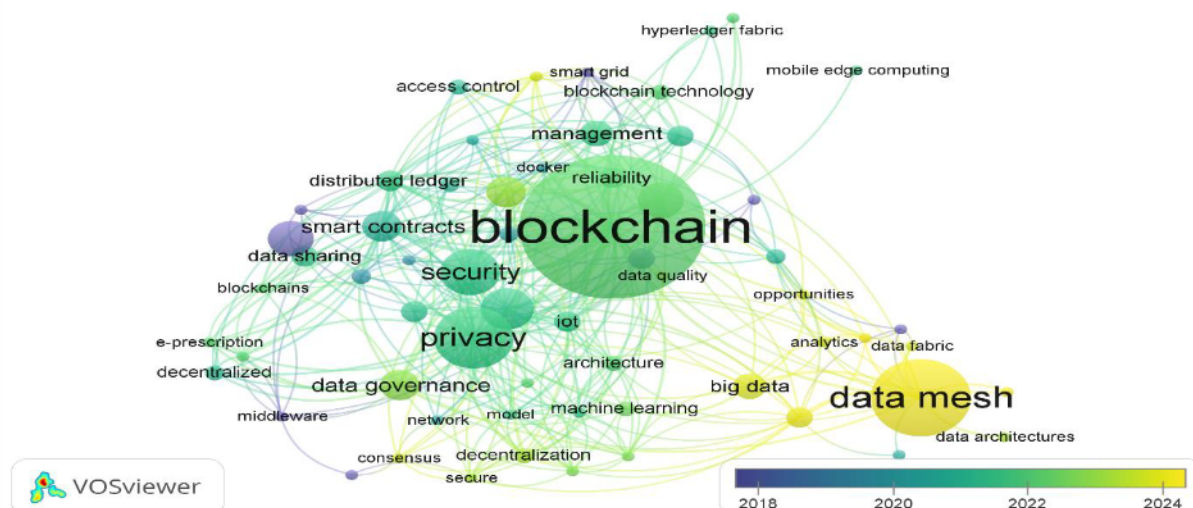


Figure 6. Keyword Clouds in the WOS Database

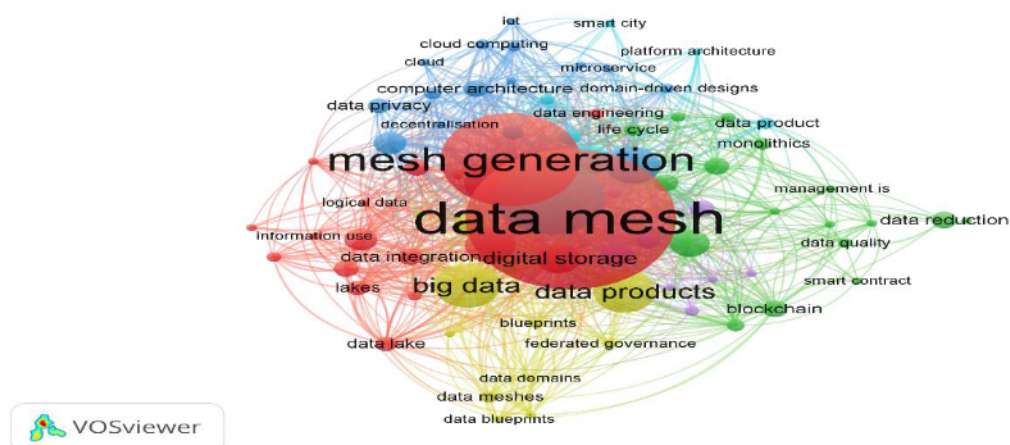


Figure 7. Co-Occurrence networks of keywords in the Scopus database

TABLE VII. TOP KEYWORDS AND OCCURRENCES

Keyword A	Keyword B	Co-Occurrence Frequency	Cluster	Relation Type
Data Mesh	Metadata	18	1	Governance foundation
Blockchain	Smart Contracts	14	2	Enforcement mechanism
Data Mesh	Ownership	13	1	Core principle
Self-Serve Platform	Scalability	12	3	Technical scalability
Interoperability	Federation	11	4	Multi-domain governance
Blockchain	Privacy	10	2	Security enablement
Metadata	Data Catalog	10	4	Discoverability & reuse
Data Mesh	Automation	9	3	Workflow efficiency

D. Co-Citation Network of Sources

Among the 26 sources that published articles, 4 clusters have emerged based on the number of citations to each other. The largest cluster is related to engineering journals, centered around the journal "*Lecture Note in Computer Sciences*". Subsequent clusters include journals related to the fields of computer science and management. Figure 5 shows the co-citation network of authors in the "Web of Science" database, where the unit of analysis is the number of cited articles, and the minimum number of citations for an article is 3. For each of the 52 cited references in 5 clusters, the total strength of the co-citation links with other cited references will be calculated. The cited references with the greatest total link strength will be selected. Table 6 shows the cluster information with thematic focus and corresponding contributions.

Figure 6 shows co-occurrence analysis when unit of analysis is author keyword corresponding to keyword clouds in the WOS database. Cluster 1 encompasses about 30% of the articles reviewed in this study. The topics of the articles published in this cluster include, respectively, the article introducing data mesh architecture [6, 7] as the cluster center, a systematic gray literature review [2] and [4] using a conceptual model of key concepts in data mesh architecture and various perspectives to design a decentralized governance model. Clusters 2 have 26% articles, with the notable and more frequently cited articles being the domain-oriented design [20] in Cluster 4 and [13] in Cluster 3. Three papers use quantitative research methods, four use qualitative research methods, and the majority of the articles published in the four clusters based on the kind of research method are library and review studies.

In Clusters 1 and 2, the subjects center on outlining the ideas and tenets of data mesh architecture as well

as the difficulties, benefits, and drawbacks of putting it into practice in businesses. Additionally, the discussion and review of various data architectures such as data lakes and data warehouses in big data governance are covered. In Clusters 3 and 4, researchers have focused on issues like the research challenges of data mesh architecture, a systematic review of the grey literature on this architecture, and its use in improving product lifecycle management. They have also used a qualitative analysis to identify challenges, motivational indicators, and the best solutions.

E. Keyword Cloud and Cumulative Growth of Words

As shown in Figures 6 and 7 "data mesh", "big data", "blockchain" and "data governance" are the most frequently occurring terms. On the other hand, the keywords "big data", "data mesh" and "decentralized" are a few key terms that have shown the most cumulative growth over the past 10 years, indicating their significant topical share in research related to data mesh architecture. Additionally, examining the trend of changing research priorities over the past decade shows that in the literature on data mesh architecture, the focus has been on general development principles, including domain-oriented design and data product thinking. However, the details of tools and methods for implementing these principles in the development and operation of the architecture to address for developing a decentralized data governance framework are lacking. Figure 7 shows co-occurrence analysis of keywords when unit of analysis is all keywords.

F. Co-Occurrence Networks of Keywords

The co-occurrence analysis of keywords shown in Figures 10 indicates that terms like "data mesh," "blockchain", "big data" and "verification" each create a concurrent network. For example, the term "data mesh" frequently co-occurs with concepts such as data governance, metadata management, distributed data management, and data platforms.

Table 8 shows analysis of Code–Theme Matrix of Co-Occurrence Relations.

G. Analysis of Co-Occurrence Networks and Trends

The co-occurrence analysis plays a vital role in exploring thematic connections and understanding the intellectual framework of data mesh research. Utilizing VOSviewer, the mapping of author keywords unveils clusters of linked concepts that highlight dominant research trajectories and emerging trends. This approach relies on the frequency and spatial proximity of keywords in selected publications (table 7), providing quantitative evidence to complement insights from qualitative analysis. The co-occurrence visualizations (Figures 7) illustrate several prominent clusters characterized by dense interconnections: Cluster I includes Core Concepts of Data Mesh Architecture with keywords like "data mesh", "data as a product", "domain ownership", "federated governance", "self-serve infrastructure". This cluster encapsulates the essential components of Data Mesh as outlined by Dehghani [6,7]. The frequent appearance of "data as a product" and "federated governance"

signals a transition from technology-focused approaches to frameworks that prioritize product-oriented and governance-conscious architecture.

Cluster II includes keywords like "blockchain", "smart contracts", "NFTs", "privacy", "auditability". This grouping centers on blockchain-driven methods for enhancing transparency, traceability, and enforcement within distributed governance frameworks. It signifies growing interest in programmable governance, emphasizing mechanisms for automation and trust. Cluster III(Data Integration and Management) includes keywords like "metadata", "data catalog", "interoperability", "data lineage". This cluster highlights key enablers for decentralized data architectures. The frequent association of "metadata" with "interoperability" underscores the need for standardized schemas and cataloging systems to facilitate cross-domain data harmonization [4, 69].

Cluster IV includes keywords like "data governance", "accountability", "compliance", "data quality". This cluster reflects themes rooted in traditional data governance models, such as those proposed by DGI and DAMA, with adaptations tailored to federated control environments. Keywords often co-occur with technical tools like APIs, automated policy frameworks, and compliance mechanisms [2, 17, 64]. By setting a minimum threshold of two occurrences, both widely discussed core ideas and emerging specialized topics were effectively captured. Keywords were organized into clusters through association strength normalization, which measures the level of co-occurrence between terms. Data mesh architecture, big data, data product, and blockchain are the most frequently occurring terms that co-occur more often with each other. These terms have shown the most cumulative growth in the past six years, highlighting their importance in research related to this architecture [5, 21]. The following sections first examine the relationship between these concepts centered on the intended architecture and then describe the research topics' trends regarding data governance based on this architecture.

TABLE VIII. TOP KEYWORDS AND OCCURRENCES

Top 10 Keywords	Total Occurrences
Data Mesh	67
Decentralized Governance	53
Big data	48
Blockchain	48
Federated Governance	41
Verification	39
Metadata Management	39
Data Ownership	37
Self-Serve Data Platform	34
Data Product Lifecycle	31
Interoperability	29
Smart Contracts	27

a) Impact of Data Mesh Architecture with Other Co-Occurrence Indicators: This architecture is built on four principles: data as a product, decentralized data ownership based on domains, self-serve data platform, and federated computational governance [6]. Data assets in an organization are presented as products that allow data consumers to discover and use domain data efficiently and independently. The domain-oriented design principle of the data mesh architecture is similar to other domain-based decentralized data architectures, using aspects like local expertise and scalability to address ownership and accountability challenges. The self-serve data infrastructure principle, as a platform, emphasizes using common tools and standards across domains while managing data teams, supporting synergy, and effective governance. The federated computational governance principle means that governance policies are implemented to respect domain autonomy while enforcing overall organizational standards. Additionally, in the context of data discovery and metadata management, it enables different teams within the organization to understand and collaborate on data products more easily [6, 10].

b) Impact of Blockchain Characteristics with Other Indicators: Ensuring high data security, reliability, and transparency in organizational data governance and facilitating the effective flow of production factors and trust in a digital organization using blockchain in data mesh architecture leads to developing a decentralized data governance framework. By applying meaningful approaches to data governance within and between organizations to facilitate and promote digital innovation using modern technologies, more innovative approaches to data governance in inter-organizational digital environments can be proposed [13, 47, 20, 61].

c) Impact of Data Governance Characteristics with Other Features: Data governance plays a crucial role in achieving a competitive advantage in overall business performance. Some prominent data governance frameworks and models include the DAMA-DMBOK framework, IBM framework, DAMA-I framework, and Data Governance Institute (DGI) framework [47, 55, 64]. From the perspective of the DGI framework, data governance is a structured system that defines decision-making rights and accountability through information-related processes. It is articulated using a set of established models, which detail who is authorized to perform specific actions with particular information, under defined conditions, at specific times, and using specified methods. In the framework ranking, this framework is described as highly efficient and agile [32, 38, 58]. In current centralized data governance practices, specialized domain teams rely on the central data management team for collecting, processing, and managing domain-specific data. This central team is becoming a bottleneck for unlocking the value of domain data. To address this issue, there is a need to transfer responsibility for domain data from the central team to specialized domain teams [7].

To further this analysis, a co-occurrence strength matrix was constructed (Figure 9). Some key insights include 1-Robust correlation between "data literacy" and "ownership & decision rights" [37], highlighting the importance of building competencies within domain teams. 2-A notable connection between "change management" and "control mechanisms" [36], underscoring the critical role of cultural shifts in the adoption of federated controls. 3-Strong co-occurrence of "privacy & security" with "compliance" and "blockchain," indicating that decentralized technologies are primarily being examined for their potential in regulatory compliance and auditability.

H. Application to Data Governance

Employing bibliometric methods such as co-citation analysis, keyword examination, and thematic classification, this review distills essential findings from the available research, shedding light on prevailing patterns, obstacles, and innovative approaches. Bibliometric analysis offers a quantitative perspective on the research landscape, emphasizing trends, prominent authors, and significant research clusters. However, to uncover more profound insights into the practical challenges and gaps in Data Mesh governance, a detailed qualitative analysis is essential. This qualitative approach enhances the bibliometric findings by examining the thematic development of research, assessing governance models, and pinpointing unresolved challenges in applying Data Mesh principles. By integrating themes from existing literature, this section seeks to present a cohesive understanding of the key governance hurdles and potential solutions in the adoption of Data Mesh.

Using a grounded theory approach, we analyzed core references to uncover key challenges, motivations, and enabling factors in Data Mesh governance. During the open coding phase, recurring terms such as accountability, domain ownership, and metadata were identified and subsequently grouped into broader axial themes like governance, data literacy, and privacy. MAXQDA-style visualizations, including category trees, code-theme matrices, and co-occurrence heatmaps, were created to systematically and visually depict these relationships.

The bar chart (Figure 8) illustrates how often each identified theme appears across the referenced studies. It emphasizes the themes most frequently addressed in the literature, such as Governance & Compliance and Change Management, which are pivotal to implementing Data Mesh. Table 9 shows the frequency of number of references corresponding to special theme.

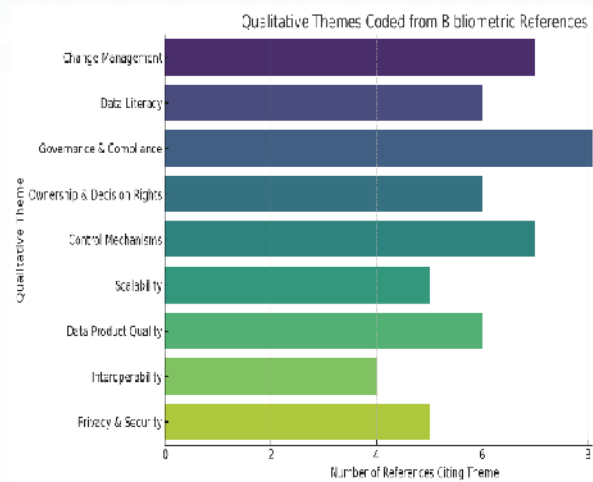


Figure 8. Number of References Citing Qualitative themes

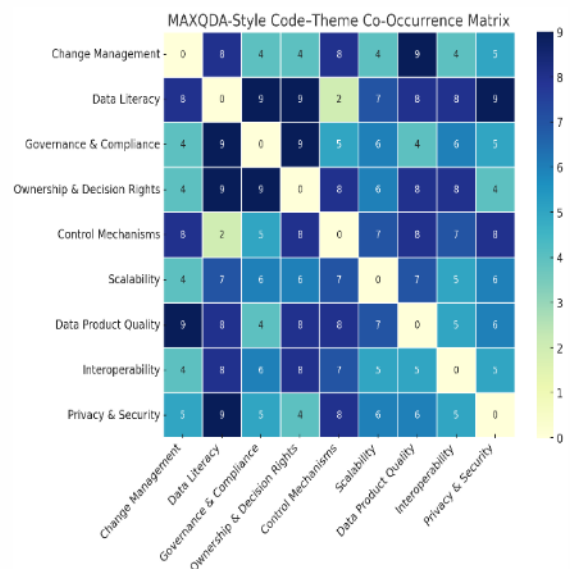


Figure 9. Frequency of co-occurrence of different themes

TABLE IX. FREQUENCY OF NUMBER OF REFERENCES

Theme	Number of References	Key References
Change Management	7	[2, 5, 19, 23, 29]
Data Literacy	6	[2, 5, 17, 19, 29, 69]
Governance & Compliance	9	[2, 4, 13, 64, 58, 59, 66, 67, 68]
Ownership & Decision Rights	6	[4, 5, 6, 7, 64, 66]
Control Mechanisms	7	[4, 13, 64, 58, 66, 67, 68]
Scalability	5	[5, 11, 22, 29]
Data Product Quality	6	[2, 4, 11, 17, 63, 64]
Interoperability	4	[2, 4, 5, 11]
Privacy & Security	5	[11, 13, 64, 66, 67]

The Figure 9 showcases the frequency of co-occurrence between various themes within the literature. Darker cells signify stronger connections between themes, as determined by co-coding in the source documents. For instance: Data Literacy is closely linked to Ownership & Decision Rights and Governance & Compliance (co-occurrence value = 9), highlighting their interconnected significance in implementing Data Mesh. Change Management's

association with Control Mechanisms (co-occurrence value = 8) suggests that the introduction of governance frameworks often brings organizational change into focus.

To answer the question of determining the key and influential indicators of data mesh architecture, the main factors a company should consider before implementing data mesh are grouped into four categories: motivation, company structure, company culture, IT history, and IT structure. To tackle the real-world practical challenges of implementing Data Mesh governance, a qualitative review of key studies was performed. This research adopts qualitative analysis to enhance the bibliometric results by examining the progression of research themes, analyzing governance structures, and highlighting lingering challenges in the application of Data Mesh principles. Each research question has been addressed to explore the role of data mesh in developing a decentralized data governance. The findings confirm the following:

- Data domains play a key role in enhancing data governance by establishing decentralized data ownership within the business. This approach empowers individuals with a deep understanding of the data to take responsibility for its management. By transferring data ownership to the business, users can directly engage with knowledgeable data owners, fostering more effective collaboration.
- Viewing data as a product enhances data governance by emphasizing its usability features. Recognizing these characteristics helps establish data as a valuable asset within the organization, leading to better management and utilization.
 - A self-serve data platform develop a decentralized data governance by implementing automated policies that establish key standards, including security, compliance, and integration. This automation promotes consistency and efficiency in managing and governing data within an organization.

The findings validate that federated computational governance significantly supports data governance. By decentralizing decision-making, it fosters a shared foundation of accountability and responsibility within the organization. Notably, IT stakeholders identified substantially more potential contributions compared to their counterparts from other departments.

Developing and implementing a data mesh comes with various challenges. These challenges arise within both the organizational and technical aspects of a company, as a data mesh is a socio-technical framework that impacts both areas. In addition, there are also certain limitations associated with a data mesh, which might act as obstacles for organizations contemplating the transformation of their data

ecosystem into a data mesh. Implementing a data mesh necessitates effective change management, as addressing resistance to change is a widely recognized challenge in the literature[2,5,17]. Additionally, adopting a data mesh requires altering both how data is approached and utilized, with the establishment of a shared understanding of its principles emerging as another challenge[17, 69]. Furthermore, a data mesh influences its users[5,17], necessitating employee training to adjust to new responsibilities. This introduces a limitation, as a data mesh demands specific skillsets within each domain and a sufficient level of data literacy [22,29].

A data mesh not only presents organizational challenges but also introduces technical difficulties. One major issue is its impact on the existing data architecture[5]. Selecting the appropriate combination of systems and engines[22] to support the data mesh proves to be a complex process, raising concerns about compatibility with the current infrastructure[5] and uncertainties in integrating tools[68,69]. Additionally, implementing a data mesh imposes pressure on the organization's network. These combined challenges make the process of adopting a data mesh highly intricate [22,29]

A data mesh may not be the right fit for every organization. Since each organization is unique, it is essential to thoroughly evaluate the motivational factors before deciding to transition to a data mesh. [6] identified six key motivational factors encouraging organizations to adopt a data mesh as follow.

- 1) Reducing bottlenecks: The central data team often struggles to process business data requests efficiently. Addressing this issue enhances time-to-market and improves the scalability of data use cases.
- 2) Utilizing domain knowledge: Delegating data responsibilities back to specific domains enhances data quality, as employees with domain expertise take charge of providing accurate domain data.
- 3) Breaking down silos: Through a self-serve platform, business units can access required data via a data catalog without needing to coordinate with other domains, thereby removing siloed barriers.
- 4) Establishing data ownership: Reassigning ownership to domains fosters a strong sense of accountability for the data created, leading to higher data quality, as poor-quality data would negatively reflect on the domain.
- 5) Adopt modern architecture: Some organizations are motivated to explore data mesh because others in their industry are adopting it. However, it is crucial to recognize that data mesh is not a universal solution, and organizations should avoid adopting it for misguided reasons.
 - 6) Reduce redundancies: Limited transparency and communication, coupled with siloed business units, often result in

duplicate efforts. A data mesh addresses this issue by leveraging a self-serve platform and establishing appropriate standards through federated governance, allowing domains to reuse work created by others. Table 10 outlines the challenges and their qualitative indicators based on the analyzed literature. Each row reflects a key theme with illustrative references. We have extracted qualitative insights from highly cited papers in Data Mesh governance by selecting key studies for analysis. Then we will analyze these studies to identify practical

data mesh governance challenges for identifying key governance challenges and possible solutions in table 10. Focusing on the DGI data governance framework, as it ranks highly among data governance frameworks[54, 60], using a literature review, one of the aims of this research is to determine key and influential indicators in data mesh architecture based on dimensions shown in the table 11 correspond to some elements of this framework.

TABLE X. CHALLENGES AND CORRESPONDING SOLUTIONS

Challenge Area	Qualitative Indicators	Possible Solutions	Key References
Change Management	Resistance to decentralization, misalignment between IT and business domains	1-Thoroughly the necessity and preparedness for implementing a data mesh, and develop a comprehensive project plan 2-Focus significantly on managing and supporting people throughout the transition process.	[2,5,17, 69]
Data Literacy	Low domain-level data proficiency, uneven understanding of data mesh principles	1-Provide training for your current employees. 2-bring in external experts, alternatively consider growing your IT team	
Accountability and Responsibility	Undefined domain roles, lack of clarity in accountability frameworks	1-Define distinct roles and responsibilities clearly 2- Implement well-defined governance policies. 3- Continuously monitor and ensure compliance 4- Ensure adherence to all applicable regulations and legislation prior to initiating the transition.	[4,58,64, 66]
Ownership and Decision Rights	Absence of domain ownership models, top-down governance legacy		[4,5,6, 7]
Control Mechanisms	Weak enforcement of data policies, lack of federated standards		[4,11,13, 66]
Interoperability	Fragmentation in metadata standards, lack of standard data product interfaces		[2, 4,5, 11]
Data Product Quality	Inconsistent metadata usage, lack of discoverability or validation mechanisms	1-Establish precise quality standards 2-Mandate the integration of metadata. 3- Continuously monitor and maintain quality levels	[11,13,63]
Privacy and Security	Compliance burdens, siloed security enforcement, lack of automation	1- implementation of standardized security measure 2-Regularly enforce security protocols 3- anonymize sensitive data 4- Perform compliance checks automatically	[13, 64, 60, 66]
Effect on the current IT infrastructure	1-Assess the compatibility of essential tools and infrastructure with the existing architecture 2- Analyze network specifications and demands 3- Thoughtfully select the optimal blend of systems and engines 4- Scalability and Infrastructure(Poor tool integration, network latency issues, legacy architecture constraints)	1-Assess the compatibility of essential tools and infrastructure with the existing architecture 2- Analyze network specifications and demands 3- Thoughtfully select the optimal blend of systems and engines	[2,5,19,21,23, 29, 34]
Effort replication	Skillfully develop and utilize the self-service platform to distribute uniform services and infrastructure.	Skillfully develop and utilize the self-service platform to distribute uniform services and infrastructure.	

TABLE XI. DIMENSIONS AND CORRESPONDING INDICATORS

DGI Elements	Mapped Themes	Corresponding Challenges	Improved Indicators/Recommendations	Key References
Accountability & Responsibility	Change Management, Governance & Compliance, Cultural Resistance	Resistance to change; unclear domain accountability	1-Formal role definitions in each domain 2-Designated data product owners 3-Responsibility matrices for domain teams	[2,4,17, 58,64, 66]
Ownership & Decision Rights	Federated Ownership, Domain-Oriented Design, Local Autonomy	Centralized legacy governance; unclear decision rights	1-Autonomous domain governance charters 2-Domain representatives in federated council 3-NFT/blockchain-enabled ownership tracking	[5,6, 7, 66]

Control Mechanisms	Policy Enforcement, Compliance, Blockchain Integration	Inconsistent governance; lack of policy enforcement	1-Automated smart contracts 2-Real-time audit logging 3-Federated policy engines for quality, access, and compliance	[4,11,13, 64, 66]
Strategy & Governance Alignment	Business-IT Integration, Strategic Alignment, Goal-Oriented Governance	Misaligned strategy; lack of governance coordination	1-Joint business-IT steering committees 2-Domain-aligned KPIs 3- Decentralized governance playbooks	[2,17, 47, 60,65]
Culture & Competencies	Data Literacy, Change Leadership, Learning Culture	Low data proficiency; change resistance	1-Domain-specific training programs 2-Data literacy certification paths 3-Leadership buy-in and change champions	[5,17, 68, 69]
Enabling Technologies	Metadata Management, Self-Serve Platforms, Interoperability Standards	Infrastructure incompatibility; siloed data systems	1- Interoperable data catalogs 2-Discoverability tools 3-Domain-specific platform abstraction layers	[2, 5, 29,30]
Centralized Coordination	Product Lifecycle Monitoring, Governance Oversight, Data Quality Control	Effort duplication; lack of unified governance monitoring	1-Federated metadata services 2-Reusability registries 3-Continuous monitoring dashboards	[30,58,63, 64]

By thoroughly reviewing the research conducted on data mesh architecture using bibliometric analysis, the key topics in the research field of this architecture include standardization of data tools and products, federated computational governance, and decentralized data ownership. Data ownership and accountability within data governance frameworks encompass issues such as decision-making, data quality, data traceability, security, and privacy [2, 10,27]. Proponents of these studies emphasize the importance of these topics and consider them fundamental challenges in data governance that require detailed examination and practical solutions [2, 5]. Especially studies focusing on data-driven decision-making and data-driven organizations highlight the need to strengthen topics related to data ownership and accountability. On the other hand, some opposing studies argue that overemphasis on these topics may lead to neglecting other issues related to data governance, such as the impacts of security and privacy on the development of data governance frameworks [2,21].

V. LIMITATIONS AND PRACTICAL RECOMMENDATIONS

This study relied on data from a specific scientific database, which may not encompass all articles published in this field. Notably, articles not published in English or those indexed in smaller or regional databases could have been excluded from the analysis. Focusing solely on English-language articles risk overlooking significant research conducted in other languages. Despite the thoroughness of this systematic literature review and bibliometric analysis on data mesh, several limitations must be acknowledged, including the temporal scope of the included literature, database restrictions, article selection bias, inherent limitations of bibliometric methods, the rapid evolution of related technologies, and potential language bias.

This literature review presents a comprehensive bibliometric analysis of the evolving field of data mesh architecture. The analysis offers a valuable overview of the current state of research by identifying key contributors, influential publications, and existing research gaps, thereby providing a roadmap for future studies on data mesh. Despite significant growth in research on this architecture, gaps remain in its implementation, especially outside the technology sector. Challenges persist regarding data governance models, regulatory compliance, and data security in decentralized environments. One promising research direction is exploring the intersection of data mesh architecture with emerging technologies like blockchain, which could offer robust solutions for decentralized trust and data ownership in developing decentralized data governance frameworks [2,8].

It is recommended that organizations consider the benefits of a self-serve data platform, not only as a technological requirement for implementing data mesh but also as a standalone tool for enhancing data governance. Such a platform, equipped with technology that automates various aspects of data governance, can provide significant value to organizations managing increasing volumes of data. Another research direction could focus on the practical implementation of both data lakehouse and data mesh approaches, exploring whether organizations can benefit from these approaches simultaneously or independently. Many organizations appear to use the data lakehouse as a platform solution when implementing data mesh. Additionally, data catalogs within data mesh architecture and their application in decentralized data governance represent a promising research area, given their future significance in data discovery, sharing, and governance. Research on standardizing data mesh architecture to ensure that data products are interoperable—facilitating product integration and collaboration across domains—is also needed [2,11].

Furthermore, the data product lifecycle—encompassing stages such as identification, design, development, testing, deployment, discovery, integration, operations, monitoring, and evolution—could serve as a valuable research area. Additional topics might include existing governance issues and challenges within data mesh architecture. Research on the principles of decentralized data ownership and federated computational governance has introduced new dimensions to data governance concerns. For instance, efficiently combining data products across domain boundaries, and potentially across organizational boundaries, depends on federated governance mechanisms that ensure data product interoperability, compliance with quality-of-service and policy requirements, and alignment with the company's overall data strategy[2,63].

VI. CONCLUSION

Data mesh addresses the scalability challenges of data systems by transforming the interaction between people and technology. It is based on four key principles: data domains, data products, federated governance, and a self-serve data platform, rather than introducing entirely new technologies[6,7]. The most important findings of this study include a systematic review of data mesh architecture and an examination of how the thematic paradigm has evolved. Recently, in data governance, topics such as data ownership, accountability, and responsibility in uncertain conditions for data analytics and AI applications have emerged. Additionally, research origins have shifted from developed countries to less developed and developing countries[8,10].

To identify current research gaps and future opportunities, the reviewed literature revealed two main issues: first, a lack of skills and resources, and second, a lack of clarity in the scope and objectives of data-related projects. Many projects begin with weak objectives, resulting in unclear definitions of necessary data, users, and purpose. Regarding skills and resources, many organizations lack sufficient technical expertise due to limited understanding of available data and its effective use, leading to concentrated technical responsibility within the IT department. The second issue is the prevalent technology-centric mindset in the industry. Many organizations approach data projects with an excessive focus on technology and tools, which causes poorly planned projects that fail to consider broader organizational needs[42, 46, 62].

Using grounded coding and MAXQDA-style outputs, the study has proposed key indicators that can be directly applied to the design of decentralized data governance frameworks in IT enterprises. Regarding the challenges of data management and utilization, cultural factors, especially the lack of a data culture, are the most common challenges. In many organizations, data sharing is often limited, resulting in challenges with accessibility and usability. This is especially common in larger organizations where data is siloed within different business units, each with its specific goals. This issue stems from the absence of a

data-driven culture, itself a result of the lack of data in the organization's business strategy[2,8,33, 36]. This cultural change should be supported by an organizational strategy set by senior management[2, 49].

REFERENCES

- [1] I. A. Machado, C. A. Costa, & M. Y. Santos, Data-Driven Information Systems: The Data Mesh Paradigm Shift. Integrated Spatial Databases, 2021.
- [2] A. Goedegebuure, I. Kumara, S. Driessen, W. J. Heuvel, G. Mondieur & D. A. Tamburri, Data mesh: A systematic gray literature review. *SKILL 2022 Gesellschaft Für Informatik, Bonn*, 29 pages, 2023.
- [3] D. Joshi, S. Pratik, & M. Rao, Data Governance in Data Mesh Infrastructures: The Saxo Bank Data Governance in Data Mesh Infrastructures: The Saxo Bank Case Study Case Study, 2022.
- [4] A. Wider, S. Verma, A. Akhtar, Decentralized Data Governance as Part of a Data Mesh Platform: Concepts and Approaches. *Thoughtworks Germany, Caffamacherreihe 7*, 20355 Hamburg, Germany, IEEE, 2023.
- [5] A. I. Machado, C. Costa & M. Y. Santos, Advancing data architectures with data mesh implementations. In *Intelligent Information Systems: CAISE Forum 2022* (pp. 10–18). Springer, 2022.
- [6] Z. Dehghani, How to move beyond a monolithic data lake to a distributed data mesh. Retrieved February 2, 2023, from <https://martinfowler.com/articles/data-monolith-to-mesh.html>
- [7] Z. Dehghani, Data mesh: Delivering data-driven value at scale. O'Reilly Media, 2022.
- [8] J. Bode, N. Kühl, D. Kreuzberger, S. Hirschl & C. Holtmann, Data mesh: Best practices to avoid the data mess, 2023.
- [9] M. Pingos, P. Christodoulou, A. S. Andreou, Security and Ownership in User-Defined Data Meshes. *Algorithms*, 17, 169, 2024.
- [10] M. Schultze, A. Wider, Data Mesh in Practice, How to set up Data-Driven Organization, O'Reilly Media, Inc, 2022.
- [11] A. Dolhopolov, A. Castelltort, & A. Laurent, Exploring the Benefits of Blockchain-Powered Metadata Catalogs in Data Mesh Architecture. *International ACM Conference on Management of Emergent Digital EcoSystems*, 2023.
- [12] B. Nemanja, I. Kumara, A. T. Damian, W. v. den. Heuvel, Privacy Engineering in the Data Mesh: Towards a Decentralized Data Privacy Governance Framework. *Lecture Notes in Computer Science*, 265-276, 2024.
- [13] A. Dolhopolov, A. Castelltort & A. Laurent, Implementing federated governance in data mesh architecture. *Future Internet*, 16, 115, 2024.
- [14] R. D. Garcia, G. Ramachandran, R. Jurdak & J. Ueyama, A blockchain-based data governance with privacy and provenance: A case study for e-prescription. In *Proceedings of the 2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, 2022.
- [15] N. Borovits, I. Kumara, D. A. Tamburri, W. J. Van Den Heuvel, Privacy Engineering in the Data Mesh: Towards a Decentralized Data Privacy Governance Framework. In: Monti, F., et al. *Service-Oriented Computing – ICSOC 2023 Workshops. ICSOC 2023. Lecture Notes in Computer Science*, vol 14518. Springer, Singapore, 2024.
- [16] M. M. Kwakye, Light-weight privacy infrastructure: A blockchain-based privacy-preservation platform for data storage and query processing. PhD dissertation, Graduate Program in Computer Science, University of Calgary, 2022.
- [17] J. Bode, N. Kühl, D. Kreuzberger & S. Hirschl, Data mesh: Motivational factors, challenges, and best practices. *arXiv preprint*, 2023.
- [18] L. Randall, Data fabric vs data mesh: 3 key differences, how they help and proven benefits. <https://www.informatica.com/blogs/data-fabric-vs-data-mesh-3-key-differences-how-they-help-and-proven-benefits.html>. Accessed 18 Jan 2024.
- [19] T. Gegenhuber, J. Mair, R. Lührsén & L. Thater, Orchestrating distributed data governance in open social innovation. *Information and Organization*, 33(1), 100453, 2023.

- [20] G. D. Rodrigo, R. Gowri, J. Raja, & J. Ueyama, Blockchain-aided and Privacy-preserving Data Governance in Multistakeholder Applications. *IEEE Transactions on Network and Service Management*, 19(4), pp. 3781- 3793, 2022.
- [21] M. Falconi & P. Plebani, Adopting data mesh principles to boost data sharing for clinical trials. In *2023 IEEE International Conference on Digital Health (ICDH)* (pp. 298–306). IEEE, 2023.
- [22] J. Divya, P. Sheetal & P. R. Madhu, Data Governance in Data Mesh Infrastructures: The Saxo Bank Case Study. The 21st International Conference on Electronic Business, 599–604, 2021. <https://aisel.aisnet.org/iceb2021/52>
- [23] A. Dibouliya & V. Jotwani, Review on data mesh architecture and its impact. *Journal of Harbin Engineering University*, 44(7), 2353–2363, 2023.
- [24] R. Winter, T. Hackl, Data mesh at scale. Exploration of current practices in large organizations. University of St. Gallen, Institute of Information Management, St. Gallen, 2023.
- [25] S. Dahdal, F. Poltronieri, M. Tortonesi, C. Stefanelli & N. Suri, A Data Mesh Approach for Enabling Data-Centric Applications at the Tactical Edge. 2023 International Conference on Military Communications and Information Systems (ICMCIS), 1–9, 2023. <https://doi.org/10.1109/ICMCIS59922.2023.10253568>
- [26] C. Voß, Identifying Alternatives and Deciding Factors for a Data Mesh Architecture. SKILL 2022 Gesellschaft Für Informatik, Bonn, 93–99, 2022.
- [27] Y. Hooshmand, J. Resch, P. Wischniewski, P. Patil, From a Monolithic PLM Landscape to a Federated Domain and Data Mesh. *Proceedings of the Design Society*. 2022;2:713-722, 2022.
- [28] A. Loukiala, J. P. Joutsenlahti, M. Raatikainen, T. Mikkonen & T. Lehtonen, Migrating from a centralized data warehouse to a decentralized data platform architecture. In *Product-Focused Software Process Improvement: 22nd International Conference, PROFES*, pp. 36–48. 2021.
- [29] T. Kraska, T. Li, S. Madden, M. Markakis, A. Ngom, Z. Wu, & G. X. Yu, Check Out the Big Brain on BRAD: Simplifying Cloud Data Processing with Learned Automated Data Meshes. *Proc. VLDB Endow.*, 16(11), 3293–3301, 2023. <https://doi.org/10.14778/3611479.3611526>
- [30] A. Almaslukh, A. Alameer, H. Alsaleh, F. Alkadyan, N. Allheeb, A. Alhadlag, & Y. Alabdulkarim, Data Mesh Meets Blockchain. *Int. J. Comput. Intell. Syst.*, 17, 27, 2024.
- [31] S. Abimbola, L. Baatiema & M. Bigdeli, The impacts of decentralization on health system equity, efficiency, and resilience: A realist synthesis of the evidence. *Health Policy and Planning*, 34, 605–617, 2019.
- [32] R. Abraham, J. Schneider & J. von Brocke, Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438, 2019.
- [33] M. Al-Ruithe, E. Benkhelifa & K. Hameed. A systematic literature review of data governance & cloud data governance. *Personal and Ubiquitous Computing*, 23, 839–859, 2019.
- [34] E. Androulaki, Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains, EuroSys '18: Proceedings of the Thirteenth EuroSys Conference, Article No.: 30, Pages 1 – 15, 2018.
- [35] M. Aria & C. Cuccurullo, Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975, 2017.
- [36] S. Black, M. Davern, B. S. Maynard & H. Nasser, Data governance and the secondary use of data: The board influence. *Information and Organization*, 33(2), 100447, 2023.
- [37] S. Bonyadi, H. Hariri, S. M. Taheri, R. Poornaghi, Presentation of a Data Quality Management Model for Data Governance Using Meta-Synthesis Method, *Journal of Information Processing and Management*, 2022.
- [38] E. Davidson, L. Wessel & S. J. Winter, Future directions for scholarship on data governance, digital innovation, and grand challenges. *Information and Organization*, 33(1), 100454, 2023.
- [39] E. Evans, Domain-driven design: Tackling complexity in the heart of software. Addison-Wesley, 2003.
- [40] F. Fathollahzadeh, N. Hariri, A. Sajadi Nejad & F. Babol-Hawaeji, R. Poornaghi, Designing a data governance model for scientific document centers. *Journal of Information Processing and Management*, 2023.
- [41] H. Ghafari, B. Kord, N. Yaghoubi, A. Koshtegar, Identifying the Antecedents and Consequences of Implementing Big Data Governance in Public Organizations Using a Mixed Methodology, *Public Management*, 2021, Volume 13, Issue 4, 2021.
- [42] K. Hendriks, *Data governance structures in data mesh architectures*. Master's thesis, Department of Business Information Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, 2023.
- [43] L. S. Jarvenpaa, & A. Essen, Data sustainability: Data governance in data infrastructures across technological and human generations. *Information and Organization*, 33(1), 100449, 2023.
- [44] J. Korhonen, Melleri, I., & M. Helenius, Designing data governance structure: An organizational perspective. *GSTF Journal on Computing*, 2. 2014.
- [45] K. Kubacki, S. Thiele, V. Lahtinen & V. Parkinson, A systematic review assessing the extent of social marketing principle use in interventions targeting children (2000–2014). *Young Consumers*, 16(2), 141–158, 2015.
- [46] T. Kunnen, Data governance in advanced analytics: Opportunities and challenges. Master's thesis, Radboud University Nijmegen, 2022.
- [47] Z. Li, F. Liang & H. Hu, Blockchain-based and value-driven enterprise data governance: A collaborative framework. *Sustainability*, 15(11), 8578, 2023.
- [48] I. A. Machado, C. A. Costa, & M. Y. Santos, Data-Driven Information Systems: The Data Mesh Paradigm Shift. *Integrated Spatial Databases*, 2021.
- [49] F. M. Majeed, N. M. Dailey, R. Khan, & A. Yunpan, Pre-caching: A proactive scheme for caching video traffic in named data mesh networks. *Journal of Network and Computer Applications*, 87, 116–130. 2017.
- [50] N. Naghshineh, F. Fahimnia, H. Ahmadian, Providing a Framework for National Data with a Focus on Data Governance Development, *Iranian Journal of Information and Communication Technology*, Thirteenth Year, Issues 94 and 95, 2021.
- [51] A. A. Neff, M. Schosser, S. Zelt, F. Uebernickel & W. Brenner, Explicating performance impacts of IT governance and data governance in multi-business organizations. *ACIS 2013 Proceedings*, 124, 2013.
- [52] C. Nwabude, C. Begg, G. McRobbie, Data governance in small businesses: Why small business frameworks should be different. *International Journal of Information Management*, 82, 101–107, 2014.
- [53] O. Olaitan, M. Herselman, N. Wayi, Taxonomy of literature to justify data governance as a prerequisite for information governance. *Proceedings of the 28th Annual Conference of the Southern African Institute of Management Scientists (SAIMS)*, 586-605, 2016.
- [54] B. Otto, Data governance. *Business & Information Systems Engineering*, 3, 241–244, 2011.
- [55] H. N. Prasetyo, K. Surendro, Designing a data governance model based on soft system methodology (SSM) in organization. *Journal of Theoretical and Applied Information Technology*, 78(1), 46–52, 2025.
- [56] M. Rodriguez, A. L. Dixon, J. W. Peltier, A review of the interactive marketing literature in the context of personal selling and sales management. *Journal of Research in Interactive Marketing*, 8 (4), 294 – 308, 2014.
- [57] D. M. Rousseau, *The Oxford handbook of evidence-based management*. Oxford University Press, 2012.
- [58] E. B. Sifah, H. Xia, C. N. A. Cobblah, Q. Xia, J. Gao, X. Du, BEMPAS: A Decentralized Employee Performance Assessment System Based on Blockchain for Smart City Governance. *IEEE Access*, 2020.
- [59] S. P. Singh, S. Gupta, Ethical data governance: An integrated approach to trust and data privacy. *Journal of Strategic Marketing*, 28(1), 77–90, 2020.

- [60] O. Solà-Morales, K. Sigurðardóttir, R. Akehurst, M. A. Linda, J. Mestre-Ferrandiz, D. Cunningham, d. P. Gérard, Data Governance for Real-World Data Management: A Proposal for a Checklist to Support Decision Making. *Value in Health*, 26(4), 32-42, 2023.
- [61] G. Vial, Data governance and digital innovation: A translational account of practitioner issues for IS research. *Information and Organization*, 33(1), 100450, 2023.
- [62] d. I. Wildt, Framing the implementation of data governance platforms, Master thesis, Business Informatics Department of Information and Computing Sciences Utrecht University, 2022.
- [63] S. U. Lee, L. Zhu, R. Jeffery, Data Governance for Platform Ecosystems: Critical Factors and the State of Practice. *arXiv Preprint*, arXiv:1705.03509, 2017.
- [64] S. U. Lee, Data Governance for Platform Ecosystems, A THESIS IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY, SCHOOL OF COMPUTER SCIENCE AND ENGINEERING FACULTY OF ENGINEERING, UNSW, Sydney, Australia, 2019.
- [65] S. Richardson, B. Walter, G. A. D. Schiava, P. Laosunthara, A. Ali & M. F. Castro, Databricks Lakehouse and Data Mesh, Part 1. Databricks. Accessed 14 June, 2022.
- [66] G. Kondova, R. Barba, Governance of decentralized autonomous organizations. *Journal of Modern Accounting and Auditing*, 15(8):406-11, 2019.
- [67] I. Kumara, S. Driessen, T. van Eijk, D. Di Nucci, D. A. Tamburri and W. J. van den Heuvel, Data Mesh Architecture: From Theory to Practice, 2024 *IEEE 21st International Conference on Software Architecture Companion (ICSA-C)*, Hyderabad, India, pp. 375-376, 2024.
- [68] M. Krystek, C. Mazurek, M. Morzy & J. Pukacki, Introducing Data Mesh Paradigm for Smart City Platforms Design. The 56th Hawaii International Conference on System Sciences, 6885-6892, 2023. <https://hdl.handle.net/10125/103468>
- [69] S. Hokkanen, Utilization of Data Mesh Framework as a Part of Organization's Data Management [Master's Thesis, University of Eastern Finland], 2021. <http://urn.fi/urn:nbn:fi:uef-20211359>.



Seyed Hassan Mousavi received his M.Sc. degree in Applied Mathematics (Computer Sciences) from the Isfahan University of Technology, Isfahan, Iran. Since 2021, he has been a Ph.D. candidate in IT management in the Department of Management at

Islamic Azad University, Najafabad Branch. His research interests include: Data Governance, Data Security, Digital Transformation.



Akbar Nabiollahi received his B.Sc. degree in Software Engineering from Isfahan University of Technology and his M.Sc. degree in Software Engineering from Islamic Azad University, Najafabad Branch. He also received a Ph.D. in

Computer Science with a specialization in Information Technology from the University of Technology Malaysia. As a faculty member at the Islamic Azad University, Najafabad Branch, he has more than 15 years of academic experience. His research interests include IT management and governance, enterprise architecture, applications of artificial intelligence in software engineering, and business intelligence.



Rasoul Asgarpour received his B.Sc in Hardware computer engineering in Islamic Azad university of Najafabad in 2006, and M.Sc. in Master of Business Administration from Multimedia University in 2010, respectively. He also received the Ph.D. degree

in Business Administration in UTM university in 2015. He is a faculty member of Department of management, Na.C., Islamic Azad University, Najafabad, Iran since 2016. His research interests include: Multimedia Marketing or Digital Marketing, Entrepreneurship, Principles of negotiations, Brand and advertising management and Change management.



Shervan Fekri-Ershad received the M.Sc. in artificial intelligence from Shiraz university in 2012 and Ph.D. in Computer engineering majored artificial intelligence from Shiraz University in 2015, respectively. He is assistant professor and a faculty member of Najafabad branch, Islamic azad university since 2015. His research interests include: Pattern recognition, Image processing, computer vision and deep learning.



Abbas Zamani received his B.Sc. degree in Business Administration from Payam Noor University in 2008. He earned his M.Sc. degree in Cultural Management from Islamic Azad University, Isfahan Branch in 2015, and his Ph.D. from the same university in 2011. Since 2013, he has been a faculty

member in the Department of Management at Islamic Azad University, Najafabad Branch. His research interests include management and social communication.