

Article

Rethinking Data Democratization: Holistic Approaches Versus Universal Frameworks

Ena Džanko ^{1,*}, Katarina Kozina ¹, Lamia Cero ¹, Antonijo Marijić ¹ and Marko Horvat ² 

¹ Syntio, Slavenska Avenija 6, 10000 Zagreb, Croatia; katarina.kozina@syntio.net (K.K.); lamia.cero@syntio.net (L.C.); antonijo.marijic@syntio.net (A.M.)

² Department of Applied Computing, Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, 10000 Zagreb, Croatia; marko.horvat3@fer.hr

* Correspondence: ena.dzanko@syntio.net

Abstract: Data democratization (DD) is a new concept rapidly becoming a game-changer, enabling companies to innovate and maintain a competitive edge in a data-driven world. This paper explores the evolution of data accessibility, from the early days of manual record-keeping to the sophisticated data management systems of today. The evolution from transactional databases to data warehouses marked a shift toward centralized data management and specialized teams, supporting the standard principles of DD contexts such as data governance (DG), privacy, management, usability, accessibility, and literacy. This paper provides an overview of the evolution of data access, from manual record-keeping to the modern data management systems of today, focusing on the challenges related to data privacy and security, integration of legacy systems, and the cultural shift required to embrace a data-driven mindset. This paper also explores both universal and holistic approaches to DD, assessing the challenges, benefits, and possibilities of their applications. An overview of industry-specific cases is included in the paper to provide practical insights that would contribute to understanding the most effective approach to data democratization.

Keywords: data democratization; data accessibility; domain-specific; data governance; big data history; data privacy; data regulations; data-driven decision-making



Citation: Džanko, E.; Kozina, K.; Cero, L.; Marijić, A.; Horvat, M. Rethinking Data Democratization: Holistic Approaches Versus Universal Frameworks. *Electronics* **2024**, *13*, 4170. <https://doi.org/10.3390/electronics13214170>

Academic Editor: Aryya Gangopadhyay

Received: 15 September 2024

Revised: 18 October 2024

Accepted: 22 October 2024

Published: 24 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The data democratization process refers to making data accessible and easily usable by all members of an organization, regardless of their role. In theory, it breaks down the barriers that traditionally limited data access to specific departments, such as information technology (IT) or data science. By empowering a broader range of employees with data, organizations can cultivate a culture of data-driven decision-making, fostering innovation and efficiency. From another perspective, some data democratization is defined as involving users—not just experts—in finding, accessing, and sharing data, while ensuring compliance and control, aligning DD with the FAIR principles (“Findable, Accessible, Interoperable, and Reusable”), according to Labadie et al. [1]. Hyun et al. [2], however, introduce a cultural aspect: the more open and diverse a culture, the better the decisions will be, and there will be greater organizational agility. In other non-IS (information systems) fields, Treuhaft [3] implies that DD is applied in urban planning, in which communities utilize data to induce social change. Each definition underlines the aspect of making the data available to overcome critical issues regarding governance or technical barriers.

A range of methodologies were used to inform our findings and proposals regarding data democratization. Our approach included comprehensive research from the existing literature, which provided foundational knowledge and insights into data democratization practices. Additionally, we drew from our hands-on experiences of various projects, allowing us to apply theoretical concepts and practical proposals to real-world scenarios. This

combination ensured a robust understanding of the complexities and challenges involved in implementing data democratization across different domains.

This paper provides a comprehensive overview of the evolution of data accessibility, tracing its journey from the early days of manual record-keeping to today's sophisticated data management systems. Section 2 provides a historical summary of data democratization, explores various definitions of the concept across different domains and cases, and offers an overview of its practical implementation through technologies. Section 3 outlines the core concepts of data democratization, explores holistic approaches to implementing it, and discusses practical evaluation factors for assessing its effectiveness. Section 4 introduces our proposed methodology for addressing data democratization challenges, presenting an innovative framework for implementing and improving data democratization practices. In Section 5, real-life use cases from different domains are introduced, including the hospitality industry, retail industry, healthcare and pharma, telecommunications, and research and science, in an attempt to recognize a pattern that occurs in working with data in the context of democratization, as well as providing an insight into the DD implementation method described in the previous section. Section 6 provides an overview of the proposed innovation for applying DD. Section 7 concludes the paper by summarizing the key insights and discussing the implications of various approaches to data democratization that could be significant for future practices.

2. History of Data Democratization

This section starts with an overview of the historical development and evolution of data democratization from traditional, restrictive access to today's full availability of data. We discuss the trade-off between data availability and security, pointing out the aspects of "Inverse Proportionality", wherein the more data are available, the greater the issues about data privacy, protection, and regulatory compliance. Further, we review data democratization across domains, showing how different fields approach data democratization uniquely. From the technological perspective, this section underlines the evolutionary path of data management practices that later on allowed modern data democratization.

2.1. From a Traditional or "Tangible" Approach

Until the 20th century, information and data were only available to the elites, scholars, and those institutions that possessed huge resources [4]. Data were mainly recorded in the form of handwritten records, manuscripts, and ledgers, which were stored and maintained manually. The advent of the printing press in the 15th century, introduced by Johannes Gutenberg, provided the capability for wider dissemination of knowledge. However, the publishing or availability of the data was still at the discretion of the owners or curators of the resources [5].

The introduction of computers marked a shift in data management. Early databases stored data electronically, replacing manual record-keeping systems. For instance, businesses started using mainframe computers to process transactions and manage inventories, streamlining operations and data storage [6].

Table 1 shows an overview of data accessibility based on the time and amount of data produced. In the pre-15th century, with manual transcription producing only about 1 to 2 pages daily, access to information was limited to scholars and the elites only [7]. The static printing press increased the output to about 40–60 pages per day, while still granting very limited access [8]. With the invention of the printing machine, the amount of accumulated data drastically increased, making books cheaper and therefore more widely available [9].

Table 1. Overview of data accessibility.

Method	Data Produced/Day (Approximate)	Total Days Produced by Era	Impact on Data Accessibility and Usability
Hand Transcription (Pre-15th Century)	1–2 pages (~1–2 KB/day)	Very limited (<1000 documents/manuscripts)	Restricted to scholars and elites, limited access
Static Printing Press (9th Century China)	40–60 pages (~40–60 KB/day)	Limited (<10,000 documents)	Increased access, but still largely restricted
Gutenberg Printing Press (15th Century)	200–300 pages (~200–300 KB/day)	20 million (~20 TB)	Data became more widely accessible; cost dropped, leading to broader societal access and improved literacy rates
Early Modern Printing (19th Century)	Thousands of pages (~5–10 MB/day)	(~100–200 TB)	Mass production allowed for universal literacy and the widespread availability
Modern Digital Printing (21st Century)	Thousands of pages (~1–5 GB/day)	(~100–200 TB)	Near-universal access to digital access to digital information Data can be produced, stored, and distributed with low barriers

At the beginning of 20th century, the idea of data democratization was in its infancy. Data access remained centralized within IT departments and data specialists. Only a handful of employees within large organizations had the necessary training and tools to access and analyze data, creating a bottleneck in decision-making processes. During the 1950s and 1960s, the storage and management of data remained mainly with IT teams, whereas access to the storage and processing of data by non-technical staff was extremely limited [10]. This would continue well into the 1970s and 1980s, during which big data practices began to emerge and further reinforced centralized control over data access [11].

The 1990s marked a pivotal era for data democratization with the emergence of the internet. Suddenly, vast amounts of information were available online, and for the first time, the general public could access information globally with just a few clicks. As shown in Table 1, this shift in accessibility parallels earlier advancements in printing technologies, highlighting the exponential growth in data production and availability from the 19th century to the 21st century. Search engines like Yahoo! (1994) and Google (1998) transformed how people retrieved and interacted with data. Parallel to the internet's rise was the increasing accessibility of personal computing. The proliferation of desktop computers in offices and homes empowered a larger number of people to interact with data more directly.

The end of the 1990s also saw the emergence of the data warehouse, a central repository of data that allowed for efficient data storage and retrieval. Organizations began investing in enterprise data systems to consolidate and analyze data from various departments [12]. Though centralization was still a dominant theme, this laid the groundwork for broader access to data within companies [13]. This era also saw the emergence of business intelligence (BI) tools that provided insights from data stored in warehouses, empowering decision-makers with actionable information [14]. These BI tools became an integral part of modern data-driven strategies by providing real-time analytics and support for decision-making [15].

The early 2000s saw a change in data democratization with the advent of self-service analytics. Tableau, founded in 2003, and Qlik, founded in 1993 but gaining broad acceptance in the 2000s, ushered in this revolution [16]. These tools provided intuitive, user-friendly interfaces for users without formal training in statistics or information technology to create data visualizations and develop reports. Meanwhile, cloud computing was also accelerating the pace of democratization. The actual marking of Amazon Web Services

(AWS) in 2006 was a critical juncture during which organizations approached and managed the storage, processing, and access of data. AWS transitioned on-premises infrastructure toward scalable, flexible, and cost-effective cloud solutions, opening doors for organizations to handle vast amounts of data in an efficient way [17].

The advent of big data technologies (such as Hadoop) addressed the challenges of storing and processing big volumes of data. For example, Google adopted MapReduce to analyze large datasets distributed across clusters of commodity hardware, setting the stage for scalable data processing and analytics [18]. In Figure 1, we can see the exponential growth in data volumes over time. It is estimated that 90% of the world's data were generated in the past two years.

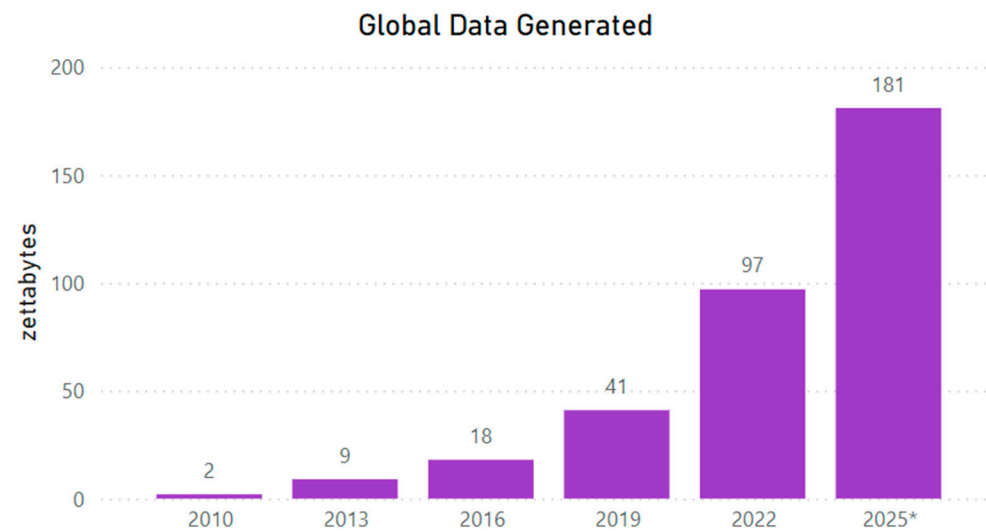


Figure 1. The graph illustrates how much data has been generated from 2010 to 2025 (*estimated). The 97 zettabytes generated in 2022 are expected to increase by over 186% in 2025, hitting 181 zettabytes [19].

The open data movement, which occurred in the mid-2000s, represents a remarkable shift toward making data freely available to the public. In addition, it has been propelled by different governments and organizations interested in enhancing transparency, firing up innovation, and fostering greater levels of citizen engagement due to the publishing of datasets openly [20]. Key initiatives related to this push were taken up by various governments in the form of data.gov in the United States and data.gov.uk in the United Kingdom, providing a single platform where a huge amount of public data should be readily available to any layperson [21].

Weerakkody et al. [20] explore the usability of open data from the perspective of a citizen. The findings indicate that open data initiatives have opened up ways for the public to access information that could be useful to achieve accountability, but at the same time, they have become more transparent. According to the authors, open data indeed has enormous potential to significantly facilitate engagement among the citizens since the provided tools and datasets mean informed participation in governance and policymaking.

Big data, artificial intelligence (AI)-driven, cloud computing and software as a service (SaaS) platforms democratized access to advanced analytics tools. Companies like AWS and Microsoft Azure offered scalable computing resources and managed services for data storage and analysis. This era also saw the rise of self-service BI tools like Tableau and Power BI, empowering business users to explore and visualize data without relying on IT departments [22].

This was a transformative period for data handling, infamously coined as the “Big Data” epoch, since it came with a real explosion of advanced technologies that increased data generation rates conspicuously. Challenges gave birth to technologies like Apache Hadoop and Apache Spark, which, for the first time, made it possible for organizations to

process and analyze huge chunks of information in a much more efficient manner. These technologies democratized data to a large extent, increasing the usability and reaching a larger class of users within the organization. Artificial intelligence (AI) and machine learning (ML) further democratized data by building predictive analytics and insights directly into cloud platforms, including Microsoft Azure and Google Cloud. Both these platforms had prebuilt models of AI and ML, which let users perform complex analysis of data, even those without much technical knowledge. The rise of cloud computing and platform as a service (PaaS) solutions was quite instrumental during this era of change [22]. PaaS provides scalable and flexible environments to develop, deploy, and manage applications, thereby increasing the reach and usability of data analytics tools. This was in addition to organizations developing data governance frameworks that balanced open data access with the need for data quality, security, and compliance. Chief data officers were very important in managing such frameworks, promoting data literacy, and making sure that data practices were aligned with organizational objectives and regulatory requirements.

The 2020s demand real-time access to data with much more advanced automation. The need for organizations to make decisions on time and with insight increasingly calls for live streams of data. Along with great power, real-time access to data and automation bring vast ethical and privacy concerns. Important regulations like the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) have brought up a very strong case for responsible data management. Organizations now face the dual task of making data more accessible and guaranteeing that these data are used responsibly and in accordance with privacy requirements. Only in this way can transparency and accountability in data treatment be addressed to regain confidence in a world where data have become more pervasive, says Zhou [23].

2.2. The “Inverse Proportionality” Perspective

In the days of paper-based data, security was implicit but usability was seriously limited. The information was physically stored, often under lock and key, which in many ways ensured confidentiality and protection against unauthorized or wider access. This way, sensitive information would only be available to those who had access to it, thus commanding high-grade security but all too often at the expense of accessibility and efficiency. This translated to slow data access that was not only error-prone but impeded decision-making and innovation.

Coveo’s report [24] found that the average employee spends 3.6 h daily looking for relevant information, and that number is even greater for an IT employee, who spends 4.2 h daily on average. Most workers report that they are slowed down or prevented from finding needed information for many reasons: multiple application storage (44%), irrelevant or outdated information on the company intranet (31%), not knowing where to look (30%) or not having permission to access the information they need (26%).

Advances in computing technology and access to the internet have made it very easy to create, share, and store a vast amount of data. Now, through the use of cloud storage and online databases, any authorized person can easily locate information and analyze it. This is where accessibility itself fosters innovation: businesses find new uses for the data in their search for competitive advantages, and researchers work together across continents to achieve scientific breakthroughs. This democratization of access to data has also raised concerns about security and privacy, as is shown in Figure 2, which shows the increasing volume of data requiring security measures. It is projected that by 2025, over 85% of data will demand some form of protection, not only personal identifiable information (PII) but also financial and medical data [25]. The digital data residing in servers and traversing over networks is prone to cyber threats in the form of hacking, data breaches, and unauthorized access.

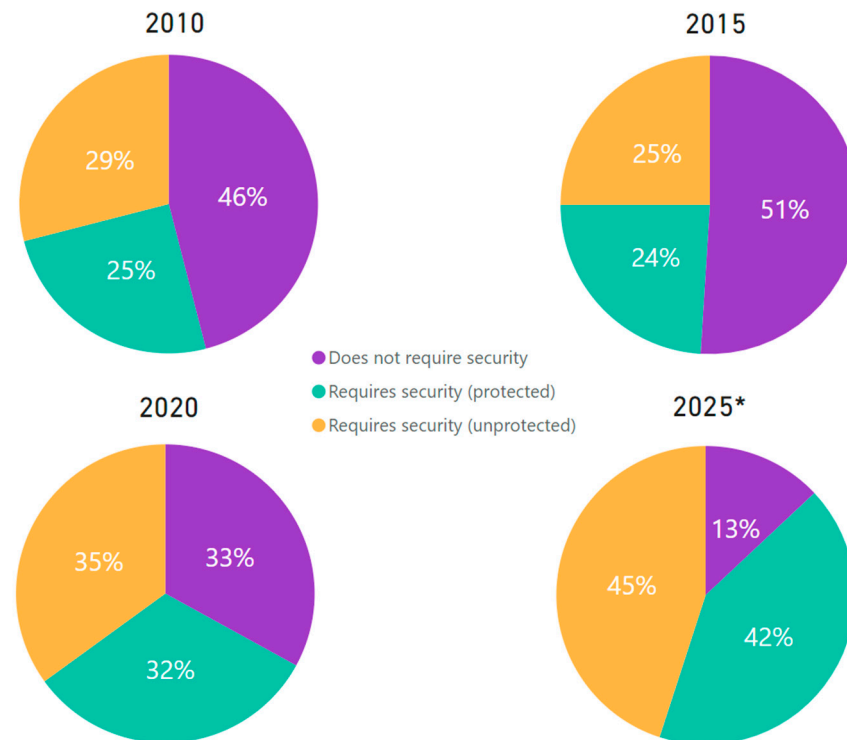


Figure 2. Volume of data which requires security measures. While most frequently accessed and used data do not require security (for example, camera phone photos, digital video streaming, open-source data, and public website content), some data significant to use cases require it, such as PII, financial data, and medical records. The percentage of data requiring security will be over 85% by 2025 (*estimated).

That is what we call a double-edged sword, meaning that where sharing the data is crucial, for instance, in medical research, we find the most caution. Therefore, data encryption, secure authentication protocols, and adherence to the regulations concerning privacy—including the GDPR—have gained much importance for organizations and governments of the world [26].

Balancing the accessibility of data with robust security measures has become a delicate act, as we arrive at the situation where security considerations slow down and prevent the progress of data democratization. Thus, the evolution from secure paper-based data storage to easily accessible but potentially vulnerable digital data exemplifies the inverse relationship in data democratization between security and usability.

2.3. Diverse Definitions and Evolution of DD Across Domains

In this subsection, various definitions and theories about DD through different domains are presented, illustrating how the concept has evolved to adjust to specific challenges as well as reflect current perspectives on this issue. DD varies significantly across different domains, reflecting distinct priorities and objectives tailored to specific contexts. In the corporate domain, the focus lies heavily on equipping employees with tools and knowledge to leverage data effectively within organizational boundaries. This approach emphasizes breaking down silos and ensuring that data are accessible to all levels of the workforce [27].

For instance, Awasthi and George [28] describe data democratization as enabling intra-organizational open data, where employees are empowered to access and analyze data independently, thus fostering a culture of data-driven decision-making. Moreover, Labadie et al. [1] highlight the importance of data catalogs in supporting data democratization within enterprises. They underscore the need for controlled yet accessible data-sharing practices aligned with FAIR principles. This perspective emphasizes the

strategic management of data access and utilization, ensuring compliance and efficiency across organizational functions.

In contrast, domains like scientific research approach data democratization with a focus on collaboration and broad data sharing beyond organizational boundaries. In medical research, for example, Bellin et al. [29] emphasize the integration of electronic medical records to facilitate data-driven insights that benefit patient care and healthcare quality goals. Here, data democratization is defined as enabling users to access and analyze data comprehensively, enhancing diagnostic capabilities and treatment outcomes.

Similarly, in urban planning, Treuhaft [3] discusses data democratization as empowering community stakeholders with Geographic Information System (GIS) tools and data to influence decisions affecting their neighborhoods. This application extends beyond traditional organizational boundaries to involve community members in leveraging data for social change and community development.

Across these domains, common threads include the promotion of data accessibility, skill development in data analysis, and fostering a collaborative culture. However, the specific nuances of data democratization vary significantly, reflecting the unique challenges and opportunities within each domain. Whether enhancing organizational agility through internal data access or empowering communities with external data resources, the overarching goal remains consistent: to leverage data effectively in order to drive informed decision-making and innovation.

2.4. A Technological Perspective: Evolution from Transactional Databases to Data Warehouses (DWHs)

We can track the evolution of DD from another perspective—based on technologies used in practice, for example—namely, how the development from transactional databases through data warehouses has laid the groundwork for the current wave of data democratization efforts.

The evolution of data management began with transactional databases, which were originally designed for the fast and efficient processing of business transactions. However, these systems were limited in their ability to support complex analytical tasks and integrate data from diverse sources, such as different stores or departments. To address these limitations, data warehouses emerged as specialized repositories optimized for data analytics. These warehouses were designed with architectures that prioritized data retrieval over data input, included pre-built reports, and accommodated massive datasets. Typically, data processing in these warehouses occurred during off-peak hours, ensuring that fresh data were available for morning reports [30].

Due to the specialized nature of data warehouse architecture and operations, which involve complex design and management tasks requiring specific expertise, organizations often centralize their data warehousing teams, sometimes referred to as BI specialists or consultants. These teams were managing data integration, ensuring data quality, and developing reports that could be understood by business users—a role often marked by a language barrier between the technical and business domains.

Data warehouses presented challenges such as tribal knowledge dependencies and the need for deep business understanding to effectively structure data for reporting and analysis. Additionally, there were two predominant approaches to data warehouse architecture: building smaller data marts for specific reporting needs or creating comprehensive big data warehouses to centralize all the organizational data.

A significant drawback of traditional data warehouses was their rigidity in handling changes. Alterations in data logic or the addition of new columns could require extensive reworking of the entire warehouse structure, impacting existing data and report integrity. Moreover, data granularity was standardized across all use cases, leading to inefficiencies such as data duplication or insufficient detail for specific analytical needs—such as detailed receipt line items versus summarized transaction data.

In summary, the evolution from transactional databases to data warehouses marked a shift toward centralized data management and specialized teams. While effective for its time, this approach posed challenges in terms of flexibility, scalability, and responsiveness to evolving business needs—an instigator of the ongoing movement toward data democratization in modern data management practices.

The journey to data democratization involves moving from a traditional to a data-driven mentality, implementing user-friendly tools, and cultivating a culture of data literacy. It is not just about technology; it is about transforming mindsets and enabling everyone to unlock the true potential of data.

3. Core Concepts and the Holistic Approach to DD

In this section, we will describe the core concepts of data democratization and its holistic approach, integrating technology, culture, education, and governance to build a data-driven organization. This section further explores how technology—from data infrastructure to governance, analytics, and AI/ML—plays its part in making data more accessible and actionable. At the end of this section, we describe the evaluation factors concerning their contribution to the success of data democratization. A pharmaceutical use case is presented to demonstrate how migrating to cloud solutions increased financial accuracy, reduced discrepancies, and promoted a data-driven decision-making culture.

3.1. A Holistic Approach to Data Democratization

The holistic data democratization approach tries to combine technology, culture, education, and governance to create an efficient data ecosystem. Technology is in charge of providing the appropriate tools and infrastructure that will enable easy access and analysis of data. Culture is relevant to creating a mindset where data are valued and used by everyone in the organization, ensuring that employees feel empowered to use data in their decision-making processes [31]. Education cultivates data literacy at all levels within the organization for the employees to develop the necessary competencies and knowledge to interpret and use data effectively. Governance ensures that data usage is regulated and secure, maintaining data quality, privacy, and compliance with relevant industry standards [32]. Together, these elements form the backbone of a successful data democratization strategy and informed decision-making across the organization.

Data democratization ensures that employees have access to the right data, with a guarantee that the information they use is relevant and accurate for their specific needs, with the aim of enabling data-driven decisions to be made. A key factor would be to enforce a data-driven culture that makes the data important and utilizes it efficiently across the organization. Therefore, this section goes over the cultural dimensions of data democratization: leadership, transparency, and collaboration.

Leadership is important in fostering a data-driven culture within organizations. Effective leaders use data as part of the decision-making process and promote environments where data are utilized at all levels. A clear vision ensures all members understand the importance of using data and maximizing its value in their daily operations [33].

Transparency in data practices builds trust across the organization. By developing infrastructures that make performance visible, organizations can improve their communication and collaboration. Regularly sharing data-driven project results strengthens trust in decision-making based on the data [34].

Collaboration between departments improves data-driven decisions and innovation. Cross-functional teams help share knowledge and ideas, embedding data-driven practices more deeply into the organization [35].

Data democratization refers to the process of enabling all members of an organization, regardless of their technical expertise, to access and analyze data. Governance in data democratization is essential to ensure data integrity and security with regulatory standards. By implementing robust governance frameworks, organizations can harness the full potential of their data assets while mitigating risks and ensuring compliance. As data continues

to play a central role in organizational decision-making, effective governance will remain a critical enabler of successful data democratization.

Data accessibility and usability are key concepts for democratization, where implementing strong and effective management systems allows access to data without compromising integrity, helping organizations unlock the full potential of their data assets [36].

Data security and privacy are important for building trust in data democratization. Compliance with regulations like the GDPR requires risk-based assessments to protect personal data while having a secure environment and ensuring responsible handling of the data [37].

Regulatory compliance in terms of data governance ensures organizations follow legal standards while protecting sensitive information. Embedding compliance into data governance practices removes the risks associated with data democratization, ensuring responsible and lawful data use across the organization [38].

Data literacy and training are crucial for successful data democratization. By investing in employee training, organizations enable employees to effectively interpret and utilize data, enhancing the overall data-driven decision-making [39]. Well-trained staff ensure that democratized data are used responsibly and efficiently.

Data stewardship and accountability are crucial for effective governance. Data stewards manage quality, integrity, and compliance, ensuring consistency across the data's whole lifecycle [40]. They address classification, security, and accuracy issues, ensuring compliance with governance policies [36]. Regular monitoring and reporting strengthen accountability and maintain reliable data [41].

3.2. *The Role of Technology in Data Democratization*

Data democratization is transforming organizational dynamics by making data more accessible and actionable for all stakeholders. In the era of big data and digital transformation, where data-driven insights are important for competitive advantages, technological advancements play a critical role [42]. Robust data infrastructure, proper data governance, a set of modern analytics tools, and AI and ML technologies can all ensure that data are accessible, usable, and valuable for any user. As highlighted in recent studies, these components are indispensable for achieving true data democratization and successfully building a data-driven organizational culture.

Data infrastructure—Proper data infrastructure will lead to effective data democratization. Common features typical of modern data infrastructure involve cloud-based platforms, data lakes, and data warehouses, which are scalable and flexible to hold an environment where data operations can be performed [42]. Adequately designed data infrastructure ensures smooth and accessible data flow; therefore, users at every level of an organization can access and utilize it with ease.

Data governance focuses on the integrity, security, and compliance of data within an organization. It involves establishing policies, procedures, and standards that govern data usage, quality, and privacy. Examples of such data governance policies may concern data classification and categorization according to sensitivity levels, allowing for proper handling and access controls. Procedures can encompass routine audits and quality checks to ensure accuracy and consistency across datasets. Standards would include encryption protocols and authentication mechanisms that ensure that data breaches and unauthorized access to the data are prevented [36]. Efficient data governance frameworks ensure data are reliable, accurate, and protected from unauthorized access. This aspect is important in building confidence in the data since it ensures that only high-quality and ethical data sources serve as the backbone for data-driven insights.

Data analytics and visualization tools have become an important turning point in the journey of raw data toward insightful information. These tools help users perform complex data analyses, find patterns, and visualize trends by building up interactive dashboards and reports. By democratizing these tools, organizations create an enabling environment where employees at all levels can engage with data, derive meaningful insights, and make

informed decisions on their own accord [43]. The ease of use of analytics platforms ensures that employees, including those without deep technical expertise, can equally contribute to the data-driven processes in an organization [42].

Machine learning and artificial intelligence—The core enabler of advanced data democratization is the integration of AI and ML technologies. These technologies can automate data analysis processes, uncover hidden patterns, and provide predictive insights that drive strategic decision-making. ML algorithms and AI models can process large volumes of data with speed and accuracy, thus delivering insights that would be impossible to achieve with manual means alone [44]. Embedding AI capabilities into data democratization efforts allows organizations to enjoy better analytical capabilities and, thereby, be innovative [43].

3.3. Evaluation Factors

Efficient data management and democratization are crucial in today's data landscape, particularly in cloud data centers (CDCs), which address challenges like energy consumption and resource allocation. The findings by Zhou et al. [45] highlight the importance of evaluation factors such as energy consumption, service level agreement (SLA) violations, and task response time. In line with the findings from the adaptive energy-aware virtual machine (VM) allocation and deployment mechanism (AFED-EF) study in the article, which highlights the importance of optimizing energy efficiency and reducing SLA violations in cloud environments, our approach to DD emphasizes the need to evaluate and address key factors like data accessibility, usability, and governance to improve organizational efficiency. By effectively directing resources and refining data strategies, proper management of data can lead to enhanced operational outcomes and long-term value, much as AFED-EF demonstrated in the case of cloud data centers.

From a data democratization perspective, the evaluation factors noted above play a role. Energy consumption plays a significant role in fostering sustainable data practices, allowing organizations to redirect resources toward enhancing democratization initiatives. Furthermore, resource allocation ensures that adequate computational power is available for processing data, thereby facilitating user access to democratized information. SLA compliance is vital for maintaining reliable data access, fostering trust among users, and encouraging broader participation in data democratization efforts. Similarly, the task response time directly impacts user satisfaction, promoting engagement with democratized data. Last, ensuring high quality of service (QoS) is fundamental, as it maintains reliable data access and empowers users to effectively utilize democratized data. Together, these factors create a cohesive framework that enhances the overall effectiveness of data democratization.

Pharmaceutical Use Case

In our recent project at a pharmaceutical company, data mappings and transformations were migrated from a traditional extract, transform, load (ETL) tool, Informatica to AWS cloud solutions. This change has been proven to be transformative. Before the migration took place, the company faced significant discrepancies in financial reports, with inaccuracies amounting to millions of dollars. These issues made it difficult to make informed decisions and plan effectively. By moving to the cloud, the organization could democratize data access and ensure that teams across the board have access to reliable, real-time information. This migration not only improved the accuracy of data but also encouraged a culture of data-driven decision-making.

Table 2 outlines the key evaluation metrics used to assess the impact of data democratization improvements in a pharmaceutical use case. The metrics emphasize how effective data democratization, while following the specific needs of the organization, leads to improved decision-making, cost efficiency, and task effectiveness, ensuring that organizations meet their SLA and comply with industry regulations.

Table 2. Evaluation metrics.

Impact Metric	Description	Pharmaceutical Use Case	Expected Impact
SLA Violations	Agreed service levels are not met	Timely integration of financial data for reports	Reduced SLA violations due to improved data processing and transformation capabilities on AWS
Task Impact	Effectiveness of tasks performed by employees	Speed of financial analysis and reporting	Enhanced task efficiency, enabling faster identification of report discrepancies
Quality of Service	Quality and reliability of service delivery	Accurate financial reporting for stakeholders	Improved service quality through data validation and error detection
Data Accessibility	Ease of access to relevant data for stakeholders	Access to financial and operational data	Accessibility leading to informed decision-making across departments
Regulatory Compliance	Adherence to industry regulations and standards	Accurate reporting for compliance audits	Enhanced compliance through more accurate financial data
Financial Accuracy	Precision of financial reports and forecasts	Reliability of financial statements	Significant reduction in financial discrepancies, improving stakeholder trust
Cost Efficiency	Reduction in costs with data management	Lower operational costs through efficient data handling	Significant savings from reduced manual error correction and improved resource allocation

In our view, we cannot adopt the perspective that there is a single, universal way to implement data democratization. Instead, we believe it is essential to consider the various components—technology, culture, education, and governance—that collectively shape a comprehensive approach to effective data democratization. While these components may seem contrary to the idea of a one-size-fits-all methodology, they can be integrated to generate a more adaptable and universal framework that suits the diverse needs of organizations.

4. Innovation: A Standardized Data Democratization Framework

This section presents an innovative approach to addressing data democratization, discussing key challenges, the data democratization maturity model (DDMM), and our proposed standardized framework to contribute to the democratization process.

4.1. Data Democratization Challenges Analysis

Achieving this ideal state of data democratization presents several challenges that are both technical and cultural. From the fragmented data landscapes to the resistance from stakeholders due to concerns over data governance, security, and accuracy, organizations often struggle to implement effective data democratization.

The key challenges include data silos, where departments or business units maintain their own data stores, limiting cross-functional collaboration and decision-making. Additionally, access control poses a significant challenge, as striking a balance between accessibility and security can be difficult, particularly when managing sensitive or regulated data. Data literacy is another critical issue, as not all users possess the necessary skills to work with data effectively, which can result in misinterpretation or underutilization of available insights. Furthermore, ensuring data quality and governance is essential to avoid flawed analysis or decisions, necessitating consistency, accuracy, and relevance across multiple systems. Outdated infrastructure or incompatible systems can also hinder seamless data integration and availability.

The proposed framework is built around addressing these key challenges by focusing on structured solutions for data integration, literacy improvement, and governance, ensuring scalable and secure data democratization.

4.2. Data Democratization Maturity Model

The data democratization maturity model (DDMM) provides a structured approach to understanding the various stages of data democratization within an organization. The stages are decided by assessing some of the key segments included in data democratization. Table 3 defines the different maturity levels for data democratization, along with the criteria of each stage. The maturity levels range from “Unaware”, where data initiatives are not organized, to “Optimized”, where data democratization is fully established and clear within processes and ownership throughout the business.

Table 3. Maturity levels with their criteria.

Maturity Level	Criteria
Unaware	Information passed down internally Individual efforts, no organization-wide awareness
Initial	Initial awareness efforts to making data discoverable and usable
Development	Organizations have a partial business glossary Terms across teams standardized to some degree
Managed	Business users have a clear understanding of the data Data are classified and tagged, standardized terminology Established business glossary Cross-team collaboration
Optimized	Data democratization established (data as a product) Owners for categories of data on the business side established Business glossary kept up to date

4.3. Overview of DD Framework

To address these challenges and implement a standardized, generally applicable approach to data democratization, we can follow the data democratization model based on the maturity and current state assessment. This model allows organizations to assess their current pain points and areas for improvement, develop a strategy, and implement data democratization progressively.

Our innovative approach to addressing DD is included in the following proposal, which has been developed through extensive practical experience and research from past and ongoing projects. This proposal is designed to be flexible, allowing for adjustments based on the specific domain, project, or organization, and is intended to contribute meaningfully to the ongoing discourse surrounding DD. As such, it represents our commitment to advancing solutions in this area, while remaining a work in progress that incorporates insights from practical applications and future planning. The model, shown in Figure 3, can be represented through three main phases, each including a set of activities providing for the end result:

1. Assessment.
2. Functional design (implementation).
3. Feedback loop or progressive implementation.

Phase 1: Data Democratization Assessment and Strategic Roadmap

In the first phase of implementing a data democratization model, assessment plays a critical role in laying the foundation for success.

Activity 1.1: Main drivers and pain points

The central activity in the assessment phase consists of the identification of the main drivers of data democratization, including data accessibility, decision-making improve-

ments, and a data-driven culture. It also addresses common pain points like fragmented data sources, poor governance, and low data literacy, clarifying the motivations and challenges the initiative must tackle. By documenting these factors, the organization outlines clarity on motivations and challenges that the data democratization initiative must address.

The deliverables of this activity include the following:

1. Comprehensive report with the main drivers.
2. Common and impactful pain points faced by the organization.
3. Proposal for the direction to steer later work.

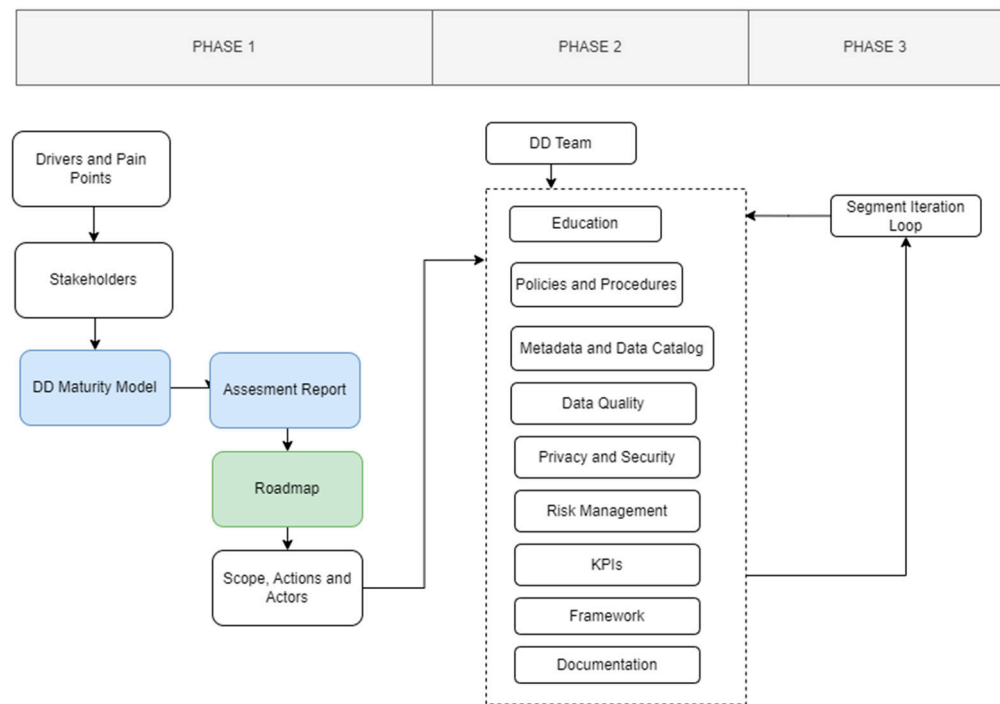


Figure 3. Illustration of the proposed methodology for implementing DD through a structured algorithm. Segmented into 3 phases: assessment, implementation, and feedback loop (scoped to each activity in the implementation phase).

Activity 1.2: Stakeholder identification

The identification of all relevant stakeholders has to be performed once the drivers and pain points have been determined. They will be departmental heads or teams directly impacted by the data democratization, IT teams, data scientists, and business analysts, among others. The next step that the organization has to take is to identify new potential stakeholders who will act as stewards in the future of this democratization process, to ensure the sustainability of the initiative in the long run. Engaging such stakeholders in the early stages of these steps ensures alignment and accountability. The result or deliverable of this phase should be formatted in an RACI matrix of stakeholders [46].

An RACI matrix is a tool used to clarify roles and responsibilities for implementing data democratization initiatives within an organization. An RACI reflects who is ultimately responsible (R), accountable (A), consulted (C), and informed (I) in the process of DD.

The RACI matrix outlined in Table 4 defines the roles and responsibilities of key stakeholders in the data democratization process, including the data steward, data democratization team, business analyst, and IT team. Each task is assigned a clear designation of who is Responsible for the execution, Accountable for oversight, Consulted for input, and Informed about progress, ensuring a structured approach to collaboration and accountability in achieving data democratization goals.

Table 4. Example of the RACI matrix.

Task	Data Steward	DD Team	Business Analyst	IT Team
Data access	R	A	C	I
Data policies	C	A	R	I
Data quality	A	R	C	I
Business data	I	C	A	R
Data processes	C	A	R	I
Education	I	R	A	C
Data usage monitor	R	C	A	I
Report on DD process	C	A	R	I

Activity 1.3: Current state assessment and DDMM

The prevailing data landscape is assessed in greater detail during this stage. This is the point at which the organization reviews the available data infrastructure, the state of data governance and democratization, technical capability, and the data access framework. An assessment of this type underlines a number of strengths, a number of weaknesses, and correspondingly, areas for improvement or change; hence, providing a starting point for data democratization. It uses a sort of maturity model to determine readiness. Because it is going to measure the organization's current level of maturity, the leadership is in a position where it can identify gaps and prioritize areas for development to help shape the strategic roadmap that will be very helpful in driving the democratization initiative.

To determine the organization's maturity level for data democratization, the following methods and activities are used:

- Data awareness review: assess the organization's overall awareness of the value, use, and accessibility of data across all levels. Mostly performed through a set of surveys.
- Data infrastructure review: Evaluate current data storage, processing, and integration capabilities.
- Technical capability assessment: analyze tools and technologies for data management and analysis.
- Data governance evaluation: assess existing governance policies, data stewardship, and compliance measures.
- Data access framework analysis: review user permissions, access ease, and data discoverability.
- Strengths, weaknesses, opportunities, and threats (SWOT) analysis: identify strengths, weaknesses, opportunities, and threats related to data democratization.

Activity 1.4: Assessment report and roadmap

The outputs of the analysis phase are given in the comprehensive assessment report. The latter defines the major drivers, pains, and the as-is and maturity status of the organization. This report is a reference document for key stakeholders and allows them to grasp the readiness status of the organization and areas to be developed concerning the data democratization perspective. Based on the report, a corresponding roadmap of activities for DD implementation is formed.

Activity 1.5: Confirm scope, key actions, and actors

With the assessment complete, the scope of the data democratization initiative is confirmed by the organization. This means the identification of specific areas of concentration, the setting of project boundaries, and confirmation that all the key players—from data stewards to executive sponsors—are in alignment and committed to the subsequent phases of the road map. This step ensures that everyone is on the same page and that the initiative has the required support moving forward.

After completing Phase 1, the deliverables include key prerequisites to start the phase of implementation:

1. Report on main drivers and pain points.

2. List of stakeholders included in the process.
3. Current state report.
4. Strategic roadmap and project scope.

Phase 2: Functional Design—Implementation

Activity 2.1: Team structure for DD

The building block of any successful DD initiative is a well-constituted and functional DG team with clear roles, such as data stewards, governance leads, and data custodians. They would manage the treatment and democratization of data. Among other responsibilities, setting standards, proper access to data, and oversight in the democratization exercise remain core activities of the team. Well-defined roles and responsibilities ensure the team manages access, usage, and compliance with data. A practical example of a DD team could include the chief data officer (CDO), who oversees the overall data strategy, data stewards, data engineers, data analysts, and data governance lead.

Activity 2.2: DD fundamentals education for stakeholders

This alignment would ensure that everyone involved, from relevant stakeholders down to end-users, understands the principles of data democratization and relates them to data governance. The basics of data democratization training would involve access policies for data, data quality standards, and ethical use of data. Accordingly, alignment from leadership to end-users can help in creating a culture within the organization toward the broader goals of democratization.

Activity 2.3: DD policies and procedures

A comprehensive policy and procedure framework is required to operate the democratization process. Such policies will dictate how data should be accessed, shared, and used within the organization. Well-drafted policies ensure that users are aware of their rights and responsibilities to handle the data the right way and the repercussions of non-compliance. Procedures will also involve how new data are onboarded into democratization, how data quality is managed, and how data usage aligns with business goals or objectives.

A standard example of points included would be the following:

1. Data access policy—defining access levels based on roles.
2. Data-sharing procedures—standardize data sharing and requests about data.
3. Data quality management policies—periodically assess data accuracy and address discrepancies.
4. Data onboarding procedures—guidelines for evaluating new data sources to ensure quality and relevance before onboarding.
5. Data usage alignment—reports from departments on data usage related to strategic business goals.

Activity 2.4: Metadata management and data catalog

Any form of democratization requires strong metadata management that can allow the data to be found and understood by all users. This means the implementation of proper data cataloging that centralizes the metadata creation, storage, and access. A well-curated data catalog provides better discoverability, context to data, and democratization of data for better decision-making. This will also include identifying what catalog platform should be utilized, what metadata standards should be defined, and how data should be tagged and organized to present to users.

Activity 2.5: Data quality management

Ensuring the quality of data is the very foundation of effective data democratization. It includes activities or measures that guarantee the organization maintains high-quality, accurate, and consistent data. This activity involves establishing rules for data validation, routine auditing of data, and specifying procedures to handle data inconsistencies. Ensuring data quality assures that the democratized data are reliable and of value for decision-making at any level of an organization.

Activity 2.6: Data privacy and security

Democratization of data surely brings critical concerns regarding data privacy and security. This involves the imposition of very strict policies to protect sensitive information while allowing users to access the needed data. Adhere to relevant data privacy regulations, such as the GDPR or the Health Insurance Portability and Accountability Act (HIPAA). In addition, it is vital to ensure that access controls are provided to prevent unauthorized access to sensitive datasets. The democratization framework should also embed regular security audits and monitoring protocols so that breaches and misuse are discouraged.

Activity 2.7: Risk management

The democratization of data brings with it a greater likelihood of mismanagement, breach, or abuse of data. Putting a risk management framework in place involves identifying possible risks, defining strategies that mitigate those risks, and developing plans of response in case those risks do occur. This activity includes compliance adherence to legal and regulatory requirements, the monitoring of data usage for abnormalities, and setting protocols in motion that quickly fix issues should they arise during data access or use.

Activity 2.8: Define DD metrics and key performance indicators

Performance indicators and metrics should be defined to measure the outcome of the data democratization effort. Such metrics can be about data access efficiency, user engagement with democratized data, and the impact of democratized data on decision-making. Clearly set key performance indicators (KPIs) enable monitoring the progress, underlining further improvements needed, and ensuring that the democratization initiative delivers real value for the business.

Following this, we could overview a practical example of some KPIs introduced in the DD scope:

1. Metric: data access efficiency—average time taken to access data across departments.
KPI: target to reduce access time by 30% within six months.
2. Metric: user engagement—number of unique users accessing the democratized data platform monthly.
KPI: aim for a 50% increase in unique users within the first quarter.
3. Metric: decision-making impact—percentage of business decisions supported by data-driven insights.
KPI: strive for 80% of decisions to be data-informed within the next year.
4. Metric: data quality improvements—frequency of data quality issues reported by users.
KPI: reduce data quality issues by 40% within the first year of implementation.

Activity 2.9: Define DD framework

Establishing a proper data governance framework aimed at data democratization will help in charting the course for data democratization. It lays down the rules and regulations regarding data access, data sharing, and usage so that democratized data across organizations is used in a very responsible and effective manner.

Activity 2.10: Documentation and communication

Well-documented logging of all processes, policies, and frameworks would contribute to keeping the democratic process of data transparent. The activity involves creating a comprehensive knowledge base in which users can obtain documentation relating to data usage, governance policies, as well as catalogs of data. Effective communication ensures that stakeholders are well-informed about updates, changes, and new procedures with regard to democratization. This would help to create a culture of collaboration and compliance.

Phase 3: Iteration and feedback loop

In this last phase of the data democratization process, the focus shifts to iteration and feedback loops, which are essential for continuous improvement. This phase begins with

collecting feedback from stakeholders regarding their experiences and challenges in using democratized data. This feedback is analyzed to identify areas needing refinement, leading to necessary adjustments and enhancements to policies, tools, and processes. Following these adjustments, organizations must reassess all the elements of the data democratization framework, including data quality, user engagement, and governance practices, to ensure alignment with the organization's goals. The key metrics and KPIs are measured again to evaluate the impact of the changes made, and hypothesis testing is employed to validate assumptions about user needs and data effectiveness. This systematic approach allows organizations to iterate on their data governance strategies and foster a culture of open communication and responsiveness, ultimately driving greater adoption and maximizing the value derived from democratized data.

4.4. Importance of Hypothesis Testing and Iterative Loops in DD Implementation

Perhaps one of the most important but commonly overlooked aspects of many data democratization efforts is the use of hypothesis testing and inferential statistics to ensure that the democratized data are high quality and can support decision-making. As debated in [47], reliance on large datasets without proper statistical assessment of the reliability of the data may result in misguided conclusions and subsequent project failures. This underlines the embedding of strong statistical procedures into the processes of data democratization to avoid such pitfalls.

Hypothesis testing thus plays an important role in proving those assumptions that are made regarding data quality, accuracy, and relevance in data democratization. The danger in democratizing data without proper testing is that such flawed or biased data could lead users to incorrect conclusions, resulting in major organizational risks. It provides a structured way of validating assumptions about the data state and quality. For example, when a company consolidates data from different departments into a unified data access layer (DAL), it might hypothesize that the combined dataset is free from inconsistencies such as duplicate records or missing fields. Without testing this assumption through statistical methods like chi-square tests for independence or t-tests to compare the means between datasets, there is a high risk of errors going unnoticed, which could lead to flawed analyses down the road.

Before going into the practical examples of hypothesis testing in DD, it is important to understand why specific tests like the chi-square and *t*-test are valuable in this context. Chi-square tests are used to analyze categorical data by comparing the observed and expected frequencies to determine if there is a significant difference between groups. In DD, this is particularly useful for evaluating data quality across multiple sources, ensuring consistency, and identifying discrepancies. On the other hand, *t*-tests are employed to compare the means of two related datasets. This is critical in data integration efforts, where we need to ensure that combining data from different systems does not introduce significant errors or distortions. These tests provide a statistical foundation for validating the quality and reliability of data before the data are democratized for wider use.

The following examples illustrate how these tests can be applied in real-world DD scenarios:

- **Data quality validation:** A financial institution merging customer transaction data from multiple branches might hypothesize that transaction error rates are consistent across all locations. By conducting a chi-square test, they can identify significant variances in the error rates, revealing branches with poor data entry practices that need improvement before broader data access.
- **Data integration testing:** The hypothesis in a healthcare organization would be that there is no duplication of records and lost information when two sets of patient records are integrated. A paired *t*-test would contribute to the checking of some key attributes, such as age and medication history, to ensure the integration process is smooth and the data are accurate for analysis by healthcare providers and researchers before democratization.

There is also a need for iteration loops when implementing data democratization. Iteration is indispensable during this process, since over time, as an organization continues to gather feedback and insight, the need for continuous refinement of data models, governance frameworks, and access mechanisms arises. This approach would enable the organization to cope with new data sources, changes in regulatory requirements, and changes in user needs to keep the process of democratization dynamic and relevant. In such iterative loops, statistical methods become very important, as they provide the mechanism to evaluate whether the changes introduced genuinely improve the quality of the data, user experience, and decision-making outcomes.

The methods of inferential statistics are also much needed when assessing the success of the DD initiative, such as by assessing user engagement with democratized data, monitoring the quality of decisions resulting from these data, and providing an overall business impact. For example, statistical tests show whether democratized data have led to better decision-making across departments or if there are areas where the data are underutilized or misunderstood. Such analyses are crucial in ensuring that the democratization effort is indeed realizing value rather than simply increasing data access without a corresponding increase in data-driven success.

In conclusion, the integration of hypothesis testing and iterative loops is crucial for a successful data democratization strategy. These methods ensure that the data being democratized is reliable and that the democratization process evolves in response to real-world challenges and user feedback.

5. Data Democratization Through Domains

Data is an asset that, in today's digitized world, leads business strategy and operations across all business domains. As a company with clients across all major business domains, we have witnessed how data democratization affects different industries.

In this section, we will first go through the universal benefits and challenges of data democratization. Then, we will go through our use cases for industries like the hospitality industry, retail industry, healthcare and pharma, telecommunications, and research and science. For each one, we will describe how data democratization supports business improvements and enhances decision-making, together with discussing special problems and challenges they face. This will show a wide understanding of data democratization regarding different fields and what impact is produced in each of them.

5.1. Universal Challenges and Benefits

Data democratization has become one of the most important strategies for organizations looking to upscale their business and those that would like to gain the most value from data in all business domains. It enables employees at every hierarchical level involved in making decisions to access and use data, thereby creating an inclusive and knowledgeable organizational environment. However, despite the significant advantages linked with the democratization of data, its implementation has significant challenges and opportunities that are not industry-specific. This section is going to elaborate on the major benefits and challenges that organizations from every industry are facing in the journey toward data democratization.

The most common challenges are the following:

- **Data privacy and security:** The more an organization opens access to data, the more critical the related privacy and security of such data becomes. From guest preferences in hospitality to patient medical records in healthcare, a large volume of personal data is exposed or vulnerable to misuse. This forces industries to operate within rigid regulatory frameworks such as the GDPR and HIPAA, or industry-specific standards that require severe measures for data protection. Finding a balance between open access and data security is challenging for any business.
- **Integration of legacy systems:** Most organizations, especially those dealing in conservative industries like health, hospitality, and telecom, still rely on legacy systems

that are not conceptualized to meet modern data ecosystem needs. Most of the legacy systems are built to operate in silos; creating seamless data flow and integrations across entities is difficult. Integration or replacement of legacy systems remains one of the most pervasive tasks, given the high costs and laboring tasks, while companies embark on a journey toward data-driven businesses.

- **Data quality and consistency:** Effective decision-making depends on the accuracy, consistency, and relevance of data. However, quality is one of the issues that must be dealt with in most organizations. Poor quality data further lead to poorly informed decisions since the information could be incomplete, outdated, or conflicting, which in turn limits the whole effectiveness of data democratization.
- **Cultural resistance:** Becoming a truly data-driven business requires a transformation not just in technology but also in organizational culture. Most industries, especially those that use gut feelings to make decisions, will encounter fierce resistance from employees who do not understand the analytics tools. Data literacy and change management programs are crucial to overcome this challenge.
- **Data governance:** Strict policies and procedures are crucial in ensuring responsible data access and usage. Data democratization raises issues of who should have access to what data for what purpose and how it will not be misused or misinterpreted. Poor data governance results in inconsistent practices, security breaches, and ethical dilemmas, especially when sensitive information is involved. More importantly, data governance requires attention from all sectors if organizations need to preserve the trust and integrity of data.

The most common benefits are the following:

- **Improved decision-making and agility:** Data democratization enables all levels of employees to make informed decisions based on current data. From the retail manager who may tweak his inventory in response to increasingly strong sales to the telecom support team who can beat customers' complaints before they start mounting, democratized data have facilitated quicker and more accurate responses. This agility allows companies to adapt faster to market changes, customer demands, and operational challenges.
- **Improved operational efficiency:** Data democratization does this through the process of optimizing the basic building blocks of business operations by presenting insights in real time. It has helped the healthcare and hospitality industries in the better management of resources to achieve an optimum level of staffing, inventories, and energy consumption. In almost all industries, data democratization allows organizations to highlight areas of inefficiencies, anticipate demands, and thereby reduce operational costs, adopting a forward-looking approach in business administration.
- **Personalized customer experience:** Access to customer data provides better-personalized services, improving satisfaction and loyalty. If information relevant to customers is available for frontline workers, an organization can offer experiences tailored to unique tastes. This turns out to be especially important in retail, hospitality, and telecommunications because such insight into the customers' behaviors and preferences may bring higher loyalty, better satisfaction, and even competitive differentiation.
- **Fostering innovation and collaboration:** Open data access encourages collaboration across departments and inspires innovation. In research-based industries like pharmaceuticals, healthcare, and scientific disciplines, it is the shared data platform that enables various teams to come together to collaborate on discoveries and treatments to bring new and innovative solutions to market. Breaking down the silo structure of data in industries allows creativity to be fostered and team experimentation with new products, services, or strategies can be performed using real-time insights.
- **Competitive advantage in a data-driven economy:** When done right, it becomes quite easy for organizations to gain a competitive advantage. Better positioned to innovate new ideas, react to changes in the marketplace, and provide improved customer experiences, this type of company is better positioned against those rivals still relying

on traditional intuition-based analytics. Companies adopting data democratization become more agile, innovative, and resilient in an increasingly data-driven landscape.

However, despite the challenges associated with data democratization, security issues, and cultural opposition being high, its benefits are revolutionary in many aspects. Data democratization is very much expected to bring about a complete transformation in the way businesses operate by improving decision-making processes, operational efficacy, and customer personalization, and by encouraging innovation. The more companies make this shift, the more they will need to deliberately balance access and governance to unlock their data's full value while maintaining trust and security.

5.2. Hospitality Industry

Data democratization has transformed the way of working in the hospitality industry. Think about a large hotel chain that used to rely on manual data entries and static reports to understand guest preferences. Due to data democratization, they implemented a centralized data platform that is accessible to employees at all levels, regardless of their technical experience. This change allowed the front desk to have immediate access to guest preferences, such as preferred room type or dining habits, which greatly enhanced the guest experience. As a specific example, one returning guest appreciated eco-friendly practices. Upon check-in, the front desk was able to recognize the preference and offered a tour of the hotel's sustainability initiatives. The guest was thrilled, and this small gesture strengthened the guest's loyalty to the brand. Another example involves a leading hotel chain that leverages data democratization to optimize its pricing strategy. Equipped with market data, hotel managers can dynamically change room rates according to local events, seasonal trends, and competitor prices. This agile pricing strategy results in significant increases in occupancy and revenue during off-season periods.

The hospitality industry must balance accessibility with data privacy. Hotels collect vast amounts of personal data, and ensuring compliance with regulations like the GDPR is critical. Additionally, many hotels still use outdated systems that do not easily integrate with modern data platforms, making it difficult to deliver a seamless guest experience across departments. Lastly, shifting from experience-based decision-making to data-driven approaches requires cultural change and comprehensive staff training.

Data democratization helps with the personalization of experiences for guests; it increases efficiency and provides a competitive advantage. Hotels can now more accurately anticipate guest needs, optimize staffing, and manage resources in real-time. All of this leads to greater guest satisfaction and loyalty. An important advantage is dynamic pricing. With access to data on markets, hotel managers are able to dynamically adjust room rates according to local events, seasonal trends, and competitor pricing. This has boosted occupancy and revenue during off-peak seasons.

5.3. Retail Industry

In the retail industry, a lot of data are present, and such data are hardly ever converted into meaningful decisions across the enterprise. This lack of sufficient alignment results in poorly defined strategic goals, deteriorating performance, and reduced attraction and retention of employees. Data can also be applied in examining consumer behavior and preferences and understanding purchasing behavior, thereby improving strategies for in-store marketing, stock management, and supply chain efficiency. For example, retailers are better placed to apply real-time sales data in optimizing their product placement or even dynamic promotional activities.

Maintaining high-quality data across diverse sources (e.g., sales data, customer behavior, and supply chain metrics) is difficult, especially when trying to ensure data security and regulatory compliance (e.g., GDPR). A centralized decision-making model prevails in retail and hence slows down the response to market changes. The data-driven culture is also built upon establishing data-literate employees and more ways of dismantling siloed systems where communication and collaboration are hampered. Data democratiza-

tion better facilitates good decision-making, improves the customer experience through personalized offers, and significantly improves operational efficiency by providing real-time insights. It helps retailers remain agile, quickly adapt to market trends, and foster innovation across departments.

The democratization of data will be at the heart of the retailers' journey to enhanced decision-making, operational efficiency, and innovation. Overcoming a host of pitfalls in ensuring data quality and security, better integration of data, and instilling a culture of empathy that is driven by data will go a long way toward unlocking data for an agile, well-informed, and collaborative enterprise that will empower critical decision-making across levels and drive substantial business enhancements.

5.4. Healthcare and Pharma: Revolutionizing Medicine with Data

Data democratization is changing the game in everything from operations and innovation to care delivery itself, from pharma to healthcare. Picture a major pharmaceutical company that once used only static reports and manual data entries to track either clinical trial progress or patient outcomes. Through data democratization, they moved to one source of truth common data platform that all stakeholders could use, independent of their technical skill level. This allowed clinical researchers to view, in real time, data regarding trial participants, such as adverse reactions or efficacy trends, thereby enhancing the accuracy and speed of the trial. A case in point: when a patient manifested a mild side effect, this could be immediately captured by the research team in adjusting the protocol of the trial to maintain the safety of the patient without sacrificing the integrity of the study. Another example might be how a healthcare professional could look toward data democratization in the optimization of managing patient care. On the one hand, providing health professionals with integrated data on patients—from medical history to real-time monitoring data—allows doctors and nurses to offer agile treatment plans, dynamically based on a patient's condition. Such agility brought about an improvement in patient outcomes, shorter stays in the hospital, and utilization of resources generally.

Major issues relate to data privacy and security, especially sensitive patient information. Regulations such as the HIPAA and GDPR need compliance. Seamless integration between legacy systems and modern platforms poses a serious problem. From the cultural viewpoint, the decisions taken by the healthcare and pharma industries have conventionally depended upon experts' intuition. Decision-making based on data would require extensive training and change at an organizational level.

Benefits accrue in the form of improved patient outcomes through appropriate treatment, operational efficiency, and speed in drug development. With data democratization, providers can optimize staffing, effectively manage resources, and deliver timely, appropriate treatments. As for pharmaceutical companies, they can speed up the delivery of drugs to the market by reducing inefficiencies during trials.

5.5. Telecommunications

Data democratization in the telecommunications industry empowers organizations to provide superior experiences for their customers and manage operations more easily. Today, telecommunication service providers provide real-time access to customer data to support teams so that issues can be resolved much faster. For example, a customer who keeps experiencing network outages may well receive proactive upgrades or personalized solutions based on his service history to prevent churn and build satisfaction. The usage data are thus provided to the marketing teams for more focused campaigns, providing the customers with exactly what they need, such as data-heavy plans for frequent streamers.

Telecom companies are faced with huge volumes of data, not just in the network being used but also in how to bill that usage. Ensuring that this is accessible while keeping regional regulations regarding privacy in mind adds much complexity to it. Integration of the legacy system with new platforms poses challenges since most telecom companies operate on systems that are anything from legacy to modern technologies. Changing the

culture into a more data-driven one across every level of the organization calls for a lot of training and investment in infrastructure.

Data democratization ensures increased customer satisfaction with personalized services and faster problem resolution. The businesses would work in an efficient manner by real-time monitoring of network performance and proper resource management. With data democratization in place, a telecom company would continue to be nimble and competitive, embracing emerging technologies like the fifth generation (5G) quickly, predicting the customer demand based on real-time data, and optimizing their services toward fulfilling them.

5.6. Research and Scientific Domain: Bridging Knowledge Gaps

Data democratization in research and science speeds up innovation and knowledge dissemination. Open access to data is changing how discoveries are made. For instance, in the field of artificial intelligence, one common dataset of labeled images allowed researchers to train new algorithms for face recognition; thus, quite substantial technological development was achieved. Democratized data have contributed a lot to developing smarter energy grids within the field of electrical engineering. Most universities and research institutions would share real-world data taken from smart meters and power grids to allow collaboration in these research efforts for efficient systems that reduce waste and costs.

Some of the various challenges include the standardization of data across diverse sources; these range from satellite pictures to the results of clinical trials. Much time is spent on cleaning and organizing data by researchers, which delays progress even more. In addition, data governance is exceedingly important, especially in sensitive fields such as biomedical research, where privacy has to be protected, and ethical standards upheld.

Democratized data fosters collaboration, speeds up the research process, and results in much stronger scientific discoveries. Shared data allow other researchers to first verify findings and then make further developments in order to solve global challenges like climate change more effectively. It is this kind of openness that drives technological and scientific breakthroughs. Table 5 shows specific challenges and benefits through different industries.

Table 5. Overview of the challenges and benefits per industry.

Industry	Challenges	Benefits
Hospitality	Privacy concerns with handling personal guest data Legacy systems that do not integrate easily with modern platforms Cultural shift toward data-driven decision-making	Personalized guest experiences Improved operational efficiency Dynamic pricing to optimize occupancy and revenue
Retail	Maintaining data quality from diverse sources Ensuring data security and compliance Cultural resistance to decentralizing decision-making	Enhanced customer experience through personalization Better inventory and supply chain management Real-time insights for agile decision-making
Healthcare and Pharma	Ensuring privacy and regulatory compliance (e.g., HIPAA) Integrating legacy systems with modern platforms Shifting from expert-driven to data-driven decisions	Improved patient outcomes with personalized care Faster drug development Operational efficiency in staffing and resource management
Telecommunications	Managing vast volumes of data from multiple sources Balancing accessibility with data privacy regulations Legacy system integration	Enhanced customer satisfaction through personalized services Real-time network performance monitoring Faster problem resolution

Table 5. Cont.

Industry	Challenges	Benefits
Research and Science	Standardizing data from diverse, often unstructured sources Data governance, especially in fields with strict ethical regulations Time-consuming data cleaning	Faster innovation through collaboration More efficient problem-solving for global challenges Accelerated discovery and validation processes

5.7. Applicability of the Universal Model Scoped to an Industry Domain

While the foundation of the DDMM remains applicable across sectors, the specific approaches and challenges can differ significantly between different domains. This section explores the unique aspects of data democratization in one example from each of two domains: hospitality and healthcare, emphasizing key differences across the stages of the maturity model.

Phase 1: Assessment

- **Hospitality:** In this sector, organizations assess their maturity by analyzing guest data, identifying key stakeholders (such as management and front-line staff), and understanding pain points related to the guest experience. For example, a hotel chain might evaluate booking patterns and customer satisfaction scores to identify areas for improvement in service delivery.
- **Healthcare:** In healthcare, the assessment phase involves evaluating patient data accessibility, identifying stakeholders such as healthcare providers and administrators, and recognizing challenges related to patient outcomes. A hospital may review its current data management practices to identify gaps in patient data collection and access that affect care quality.

Phase 2: Functional Design

- **Hospitality:** During this phase, hospitality organizations design team structures that facilitate data sharing among staff. They implement training programs focused on data literacy and develop policies governing data privacy and security for guest information. For instance, a restaurant chain might create a centralized data catalog to track customer preferences and dining trends while establishing guidelines for staff on handling these data responsibly.
KPI examples: percentage of staff trained in data literacy, customer satisfaction score (CSAT) based on feedback
- **Healthcare:** In the healthcare sector, the functional design includes establishing interdisciplinary teams that integrate data management with clinical practices. Organizations might develop detailed data governance policies, enhance metadata management systems, and implement training on data privacy regulations (e.g., HIPAA compliance). A hospital may set KPIs related to data accuracy and timeliness in patient care.
KPI examples: percentage of patient records accurately digitized, average time taken to access patient data for treatment decisions

Phase 3: Iteration and Feedback Loop

- **Hospitality:** In the hospitality domain, feedback mechanisms can be established through customer surveys and staff input, allowing for continuous improvement of data practices. For example, a hotel might regularly assess guest feedback to refine data collection methods and improve overall guest experiences.
- **Healthcare:** Healthcare organizations can leverage patient feedback and outcome data to assess the effectiveness of their data democratization efforts. By analyzing readmission rates and patient satisfaction scores, a hospital can refine its data strategies to ensure it better meets patient needs and improve health outcomes.

6. Universal Formula

In the journey toward data democratization, one of the most vital questions must be the following: Is there a single underlying formula for effective data management and access across organizations and domains, or does success depend on certain contextual factors? The idea of a single formula serving all organizations for data democratization is exciting; it is, however, elusive considering the different challenges and requirements that organizations must meet.

The universal formula concept means one set of practices and technologies that can universally serve the needs and use cases of various domains. This would involve centralizing data platforms, globalizing access controls, and setting universal data governance policies. The truth is, however, that data management and democratization are contextual activities. These could include the nature of the data itself, particular needs on the part of the users, regulatory requirements, and issues of organizational culture. Data democratization in healthcare, for example, must account for strict privacy regulations and specialized, specific data analysis tools. In retail, the approach can be much more geared toward various streams of data gathered to optimize inventory and customer experiences. What works for one simply may not work well for another, and the optimal technologies and practices thus vary.

Moreover, the evolving landscape of data management technologies and practices further complicates the search for a universal formula. Advances in cloud computing, big data analytics, and artificial intelligence continuously reshape how data are accessed and utilized, requiring adaptive strategies that can accommodate new developments and challenges.

While principles like accessibility, usability, and governance are foundational and provide the baseline for efficiency in data management, their implementation must be tuned to suit the specific domain. The universal formula may stand in theory or even in some initial steps of implementation, but focusing on the complete process requires flexibility and customization for different needs and environments. Therefore, data democratization requires a seemingly fitting approach, which must be dynamic and sensitive to context rather than rigid, following a one-size-fits-all solution.

The proposed universal formula could only emerge as a combination of established DD foundations (patterns) but with a holistic approach, integrating diverse methodologies and perspectives:

- Holistic approach: focusing on different factors affecting data management.
- Foundational principles: common and generally applicable components of DD like accessibility, usability, and core data governance.
- Combined methodologies: maturity model assessment, step-by-step implementation, hypothesis testing, iterative loops.
- Dynamic framework: customization and responsiveness to evolving data, technologies, and regulations.
- Contextual adaptation: recognizes that different domains have unique requirements and challenges, necessitating flexible implementations.

7. Conclusions

This paper highlights why data democratization, though crucial for creating an inclusive data-driven culture, faces challenges that cannot be solved with a “one-size-fits-all” approach. It traces the evolution of data accessibility from manual information capture to modern data management systems, emphasizing the domain-specific obstacles organizations face when implementing democratization strategies.

With this paper, it has been possible to provide an overview of the constant change that data democratization is experiencing and analyze the holistic approach versus the universal approach to understanding and dealing with such. Indeed, the universal approach with a one-size-fits-all formula might present general principles and basic guidelines across most contexts. Our discussion and innovation demonstrate that data democratization cannot

be approached and managed through only one valid, universally applicable strategy in most cases. Instead, the holistic approach works more appropriately, and again it should be contextualized within specific topics such as standard data concepts (data governance, ethics, literacy, lineage, management, integrity, consistency, etc.) or perspectives (cultural, educational, technical). The method acknowledges that various domains and industries have different needs that impact the way data are handled.

Several directions can be taken in the future work on data democratization. One potentially challenging area is the development of adaptive frameworks that can adjust dynamically according to the changing needs of different sectors. Research efforts could focus on developing modular and scalable data solutions tailored to the specific needs of industries such as healthcare, retail, financial services, and manufacturing. For instance, creating adaptable platforms for patient data integration, inventory management, real-time transaction processing, and supply chain optimization could ensure that data democratization practices remain relevant and effective across diverse contexts. There is also a need to research the actual level of the current and future integration of technologies, including high-level AI and machine learning, that offer new opportunities for enhancing data access and usability. This demands a more profound understanding of organizational culture and employee training implications based on successful data democratization strategies. This also involves exploring how to develop a culture of data literacy and lifelong learning that supports or permits information-based decision-making.

The proposed innovation introduces a general or core approach to addressing data democratization, offering a foundational framework that can be adjusted to specific domains. This approach emphasizes the importance of building upon standardized practices, ensuring consistency while remaining flexible enough to adapt to the unique challenges of various sectors. However, to maintain its effectiveness, this approach must be continuously refined and developed in alignment with evolving technologies, data governance standards, and organizational needs. By integrating this adaptable yet standardized model, organizations can pursue a more structured and sustainable path to achieving DD in most early phases. Addressing these areas will enhance efforts to meet the demands of an ever-changing data landscape.

Author Contributions: Conceptualization, E.D.; methodology, E.D.; validation, M.H., E.D., K.K., L.C. and A.M.; formal analysis, K.K.; investigation, L.C.; resources, E.D., K.K. and A.M.; writing—original draft preparation, E.D., K.K. and L.C.; writing—review and editing, M.H., E.D. and K.K.; visualization, L.C.; supervision, E.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Authors E.D., K.K., L.C. and A.M. are employed by the company Syntio. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Labadie, C.; Legner, C.; Eurich, M.; Fadler, M. FAIR Enough? Enhancing the Usage of Enterprise Data with Data Catalogs. In Proceedings of the 2020 IEEE 22nd Conference on Business Informatics (CBI), Antwerp, Belgium, 22–24 June 2020; pp. 201–210. [CrossRef]
2. Hyun, Y.; Kamioka, T.; Hosoya, R. Improving agility using big data analytics: The role of democratization culture. *Pac. Asia J. Assoc. Inf. Syst.* **2020**, *12*, 2.
3. Treuhaft, S. The Democratization of Data: How the Internet Is Shaping the Work of Data Intermediaries. 2006. Available online: <https://escholarship.org/uc/item/32961226> (accessed on 21 August 2024).
4. Reinking, D.; Colwell, J. A Brief History of Information Sources in the Late 20th and Early 21st Century (A simulation). 2015. Available online: https://open.clemson.edu/cgi/viewcontent.cgi?article=1065&context=eugene_pubs (accessed on 21 August 2024).
5. Cameron, F.R. *The Future of Digital Data, Heritage and Curation: In a More-than-Human World*, 1st ed.; Routledge: London, UK, 2021. [CrossRef]

6. Haigh, T. How Data Got its Base: Information Storage Software in the 1950s and 1960s. *IEEE Ann. Hist. Comput.* **2009**, *31*, 6–25. [CrossRef]
7. Reynolds, L.D.; Wilson, N.G. *Scribes and Scholars: A Guide to the Transmission of Greek and Latin Literature*; Oxford University Press: Oxford, MS, USA, 2013.
8. Febvre, L.; Martin, H.J. *The Coming of the Book: The Impact of Printing 1450–1800*; Verso: Brooklyn, NY, USA, 1997; Volume 10.
9. Eisenstein, E.L. *The Printing Press as an Agent of Change*; Cambridge University Press: Cambridge, UK, 1980; Volume 1.
10. Kerssens, N. The database ‘revolution’: The technological and cultural origins of the big-data-based mindset in American management, 1970s–1980s. *TMG–J. Media Hist.* **2018**, *21*, 7–29. [CrossRef]
11. Eze, U.F.; Etus, C.; Uzukwu, J.E. Database system concepts, implementations and organizations—A detailed survey. *Database* **2014**, *2*, J2013117.
12. Breslin, M. Data warehousing battle of the giants. *Bus. Intell. J.* **2004**, *7*, 6–20.
13. Simon, A.R.; Shaffer, S.L. *Data Warehousing and Business Intelligence for E-Commerce*; Elsevier: Amsterdam, The Netherlands, 2001.
14. Khan, R.A.; Quadri, S.M.K. Business intelligence: An integrated approach. *Bus. Intell. J.* **2012**, *5*, 64–70.
15. Marjamäki, P. Evolution and Trends of Business Intelligence Systems: A Systematic Mapping Study. 2017. Available online: <https://oulurepo.oulu.fi/handle/10024/8679> (accessed on 29 August 2024).
16. Kimball, R.; Ross, M. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*; Wiley: Hoboken, NJ, USA, 2011; ISBN 978-0471228552.
17. Marinescu, D.C. *Cloud Computing: Theory and Practice*; Morgan Kaufmann: Cambridge, MA, USA, 2018. Available online: <https://industri.fatek.unpatti.ac.id/wp-content/uploads/2019/03/212-Cloud-Computing-Theory-and-Practice-Dan-C.-Marinescu-Edisi-2-2017.pdf> (accessed on 4 October 2024).
18. Özköse, H.; Sertaç Arı, E.; Gencer, C. Yesterday, today and tomorrow of big data. *Procedia-Soc. Behav. Sci.* **2015**, *195*, 1042–1050. [CrossRef]
19. Volume of Data/information Created, Captured, copied, and Consumed Worldwide from 2010 to 2025. Available online: <https://www.statista.com/statistics/871513/worldwide-data-created/> (accessed on 3 September 2024).
20. Weerakkody, V.; Irani, Z.; Kapoor, K.; Sivarajah, U.; Dwivedi, Y.K. Open data and its usability: An empirical view from the Citizen’s perspective. *Inf. Syst. Front.* **2017**, *19*, 285–300. [CrossRef]
21. Wanchoo, L.; Lindsay, F.; Turcios, D.; Hao, Y.; Weir, H. NASA EOSDIS 20 Years of Data Usage and User Assessment in Support of Open Science Initiative. In Proceedings of the 23rd Meeting of the American Geophysical Union (AGU) (No. IN41B-0589), San Francisco, CA, USA, 11–15 December 2023.
22. Yasrab, R. Platform-as-a-service (PaaS): The next hype of cloud computing. *arXiv* **2018**. [CrossRef]
23. Zhou, L. Ethics and Data Privacy in the Era of Big Data. *J. Data Ethics* **2021**, *3*, 45–59.
24. Coveo Solutions Inc. Workplace Relevance Report 2023: The Growing Generation Gap: Is Innovative Talent at Risk? 2023. Available online: <https://ow.ly/XxVU50OSEO6> (accessed on 4 October 2024).
25. Reinsel, D.; Gantz, J.; Rydning, J. Data Age 2025: The Evolution of Data to Life-Critical. Don’t Focus on Big Data; Focus on the Data That’s Big. Seagate, April 2017. Available online: <https://www.seagate.com/files/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf> (accessed on 6 September 2024).
26. Abdullah, F.M. Privacy, security and legal challenges in big data. *Int. J. Civ. Eng. Technol.* **2018**, *9*, 1682–1690.
27. Lefebvre, H.; Legner, C.; Fadler, M. Data democratization: Toward a deeper understanding. In Proceedings of the 42nd International Conference on Information Systems (ICIS 2021), Austin, TX, USA, 12–15 December 2021.
28. Awasthi, P.; George, J.J. A case for Data Democratization. In Proceedings of the Americas Conference on Information Systems, Virtual Conference, 15–17 August 2020.
29. Bellin, E.; Fletcher, D.D.; Geberer, N.; Islam, S.; Srivastava, N. Democratizing information creation from health care data for quality improvement, research, and education—The Montefiore Medical Center Experience. *Acad. Med.* **2010**, *85*, 1362–1368. [CrossRef]
30. Vargas-Solar, G.; Zechinelli-Martini, J.L.; Espinosa-Oviedo, J.A. Big data management: What to keep from the past to face future challenges? *Data Sci. Eng.* **2017**, *2*, 328–345. [CrossRef]
31. Meng, X.L. Data Democratization: An Ecosystemic Contemplation and Coordination. 2024. Available online: <https://hdr.mitpress.mit.edu/pub/o3qm160v/release/6> (accessed on 6 September 2024).
32. Takang, T.M.; Oguejiofor Amaechi, A. Considerations for a Planned Democratizing Data Framework for Valid and Trusted Data. *J. Data Anal. Inf. Process.* **2023**, *11*, 240–261. [CrossRef]
33. McCarthy, J.; Sammon, D.; Murphy, C. Changing leadership behaviours: A journey towards a data driven culture. In Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, 5–10 June 2017; pp. 2625–2634. Available online: http://aisel.aisnet.org/ecis2017_rfp/1 (accessed on 10 September 2024).
34. Pine, K.; Mazmanian, M. Emerging Insights on Building Infrastructure for Data-Driven Transparency and Accountability of Organizations. iConference 2015 Proceedings 2015. Available online: <https://www.ideals.illinois.edu/items/73662> (accessed on 10 September 2024).
35. Samarasinghe, S.; Lokuge, S.; Snell, L. Exploring Tenets of Data Democratization. 2022. Available online: <https://arxiv.org/abs/2206.12051> (accessed on 27 August 2024).
36. Khatri, V.; Brown, C.V. Designing data governance. *Commun. ACM* **2010**, *53*, 148–152. [CrossRef]

37. Lăzăroiu, G.; Kovacova, M.; Kliestikova, J.; Kubala, P.; Valaskova, K.; Dengov, V.V. Data governance and automated individual decision-making in the digital privacy General Data Protection Regulation. *Adm. Si Manag. Public* **2018**, *31*, 132–142.
38. Ladley, J. *Data Governance: How to Design, Deploy, and Sustain an Effective Data Governance Program*; Academic Press: Cambridge, MA, USA, 2019.
39. Gupta, M.; George, J.F. Toward the development of a big data analytics capability. *Inf. Manag.* **2016**, *53*, 1049–1064. [[CrossRef](#)]
40. McGilvray, D. Data Quality Projects and Programs. In *Handbook of Data Quality: Research and Practice*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 41–73.
41. Weber, K.; Otto, B.; Österle, H. One Size Does Not Fit All—A contingency approach to data governance. *J. Data Inf. Qual.* **2009**, *1*, 1–27. [[CrossRef](#)]
42. Brynjolfsson, E.; McAfee, A. Thriving in the automated economy. *Futurist* **2012**, *46*, 27–31.
43. Provost, F.; Fawcett, T. Data science and its relationship to big data and data-driven decision making. *Big Data* **2013**, *1*, 51–59. [[CrossRef](#)]
44. Davenport, T.H.; Ronanki, R. Artificial intelligence for the real world. *Harv. Bus. Rev.* **2018**, *96*, 108–116.
45. Zhou, Z.; Shojafar, M.; Alazab, M.; Abawajy, J.; Li, F. AFED-EF: An energy-efficient VM allocation algorithm for IoT applications in a cloud data center. *IEEE Trans. Green Commun. Netw.* **2021**, *5*, 658–669. [[CrossRef](#)]
46. Ahmed, M.R. *The RACI Matrix and Its Implications: A Case of Unilever*; BRAC Business School, BRAC University: Dhaka, Bangladesh, 2018.
47. Roccetti, M.; Delnevo, G.; Casini, L.; Cappiello, G. Is bigger always better? A controversial journey to the center of machine learning design, with uses and misuses of big data for predicting water meter failures. *J. Big Data* **2019**, *6*, 70. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.