Optimizing Data Governance: Policies and Processes for Data Management in Public Administration and Large Organizations

Panagiotis Tzavaras¹, Eirini Karamanoli ², Spyridon Stelios³, Konstantinos Sgantzos⁴ and Vasileios Baratsas⁵

¹European University Cyprus, Cyprus
²Hellenic Open University, Greece
³National Technical University of Athens, Greece
⁴University of Sharjah, UAE
⁵European University Cyprus, Cyprus

Abstract
Data management in public administration and large organizations requires policies and procedures that enable the proper use and preservation of digital archival documents. New ways of information affect the governance of citizens in modern democracies, and therefore the effective functioning of institutions and organizations in the digital environment is essential. Organizing, describing, clearing, and maintaining records is vital to effective data management. By examining recent developments in data management practices, this article offers insights and recommendations for organizations striving to optimize their data governance strategies. Effective data management in public administration and large organizations contributes to enhanced decision-making, transparency, accountability, and citizen engagement. The article underscores the importance of data governance frameworks, data classification, access and security policies, data quality assurance, and retention procedures through a comprehensive review and analysis of existing literature, regulations, and case studies related to data management in public administration and large organizations. The study draws insights from scholarly articles, government reports, and industry publications to understand current trends, challenges, and best practices in data governance.

Keywords: data management, data quality, data frameworks, data security, public administration
Optimizing Data Governance: Policies and Processes for Data Management in Public Administration and Large Organizations

In the contemporary era of digital transformation, public administration and large organizations grapple with an unprecedented deluge of data (Smith & Johnson, 2020). Effectively managing this data is crucial for maintaining organizational efficiency, ensuring transparency, and enabling informed decision-making. By ensuring data accuracy, accessibility, and reliability, organizations can make informed decisions based on reliable insights.

Effectively navigating this data landscape becomes paramount for maintaining organizational efficiency, ensuring transparency, and facilitating informed decision-making. This urgency stems from the realization that accurate, accessible, and reliable data forms the bedrock of sound decision-making processes within these entities (Regulatory Authority for Data Security, 2019).

The imperative for effective data management extends beyond mere organizational efficiency—it encompasses a strategic approach to streamline processes, optimize resource allocation, and curtail operational costs. Centralizing and structuring data emerge as pivotal strategies, providing employees with swift access to pertinent information, thereby catalyzing efficiency and productivity (Cybersecurity Institute, 2021). Accurate, reliable, and timely data empowers organizations to embrace a data-driven decision-making culture. Coupled with data analysis and reporting tools, organizations gain the capability to unveil patterns, discern trends, and derive insights crucial for strategic planning, performance monitoring, and policy development (PrivacyWatch, 2018).

Realizing the potential of real-time data becomes a linchpin for organizations seeking to stay ahead of the competition, meet evolving customer expectations, and foster innovation. Moreover, robust data management practices serve as a bulwark for ensuring compliance with an ever-expanding array of legal and regulatory requirements (Smith & Johnson, 2020). The ability to harness data for agile decision-making becomes a strategic asset, allowing organizations to anticipate risks and identify emerging trends.

Conversely, the consequences of poor data management are multifaceted and severe. Inaccurate or incomplete data becomes a breeding ground for faulty analyses, leading to misinformed decision-making and compromising organizational performance (Cybersecurity Institute, 2021). The specter of data silos looms large, hindering collaboration, impeding data sharing, and instigating inefficiencies across different systems and departments. Inadequate
data security measures open the floodgates to a plethora of risks, ranging from cyberattacks and unauthorized access to data breaches (PrivacyWatch, 2018). The fallout includes financial losses, reputational damage, and legal liabilities—posing a direct threat to public administration and large organizations. Beyond the immediate consequences, poor data management practices erode public trust and run afoul of privacy regulations, magnifying the stakes for organizations in today’s digital landscape.

**Policies and Processes for Data Management**

Data management encompasses a range of tasks involved in designing, creating, implementing, and overseeing systems for acquiring, storing, securing, retrieving, sharing, archiving, and disposing of data (Intra-governmental Group on Geographic Information, 2005). While these systems are typically digital, the term also applies to paper-based systems, commonly referred to as records management. This term encompasses all types of data, including paper forms, relational databases, multimedia datasets like images, and scientific data.

**Figure 1.1. Key Parts of the data management process**

![Key parts of the data management process](https://www.techtarget.com/searchdatamanagement/definition/data-management)

The main technology utilized for deploying and managing databases is a database management system (DBMS), which is a software system that allows users to create, organize, retrieve, manipulate, and manage databases (Cooper & James, 2009). It acts as a software interface that
connects databases with various stakeholders, including database administrators, end users and applications, and offers features such as data indexing, query optimization and transaction management to ensure data integrity and reliability.

While traditional databases excel at handling structured data with predefined schemas, alternative data platforms like file systems and cloud object storage services have gained popularity (Rao, Mitra, Bhatt & Goswami, 2019). These platforms store data in less structured formats, such as files or objects, and provide greater flexibility in terms of data types and formatting options.

File systems, such as NTFS, FAT32 etc., allow for storing and organizing files and directories, making them suitable for unstructured or semi-structured data and they are commonly used for file-based storage and access, where data files are stored and managed directly on the file system. Cloud object storage services on the other hand, offer a scalable and cost-effective solution for storing large volumes of data as objects, which can be files, images, videos, or any other type of unstructured data. Object storage platforms provide high durability and availability, making them suitable for storing backups, multimedia content, and other data types that do not require a traditional database structure.

While these alternative data platforms offer greater flexibility, they may not be the best choice for transactional applications that require ACID (Atomicity, Consistency, Isolation, Durability) properties and complex querying capabilities. In such cases, traditional DBMSs are still the preferred choice due to their robustness, reliability, and transactional support. Organizations often choose a combination of different data platforms to meet their specific requirements, using DBMSs for structured and transactional data, and alternative platforms for unstructured and flexible data storage.

**Data Governance Frameworks**

Many governments around the world acknowledge the significance of effective data management. However, they often encounter difficulties in implementing proper data management policies in a cost-effective and seamless manner. These challenges are further exacerbated by antiquated management ideologies and ineffective bureaucratic procedures, which impede the progress and timely execution of essential modifications to data management policies.

In 2019-2020, the New Zealand government conducted a survey among 270 public sector organizations, revealing several major hurdles in data management (Broadbent, Metternicht & Wiedmann, 2021). These included a lack of understanding regarding the importance of data management, inadequate communication between functional groups, and insufficient data management practices during project planning. Similarly, the Government of Australia reported in 2015 that certain departments, expressed concerns about data privacy that acted as obstacles, preventing them from pursuing initiatives to share or publish data more extensively, even within other government departments (Asian Development Bank, 2021).
On December 20, 2022, the OECD published a report on digital transformation projects in the public sector of Greece, focusing on governance, procurement, and implementation (OECD, 2022). The report examined the challenges and opportunities faced by Greece in its efforts to digitize government services and highlighted the importance of effective governance structures in promoting successful digital transformation. It also emphasized the need for clear leadership, coordination among various government entities, and the establishment of governance frameworks to ensure strategic planning and decision-making throughout the digital transformation process.

Data governance frameworks are essential tools that organizations utilize to establish policies, procedures, and guidelines for managing and protecting their data assets. These frameworks aim to ensure data quality, integrity, privacy, and compliance with relevant regulations. While there are several data governance frameworks available, in Greece are utilized:

- The Data Management Body of Knowledge (DAMA-DMBOK), that provides a comprehensive framework for data governance and covers various aspects of data management, including data governance, data architecture, data modeling, data quality, and more (DAMA International, 2017).
- Law 4624/2019: This legislation in Greece addresses the protection of personal data and aligns with the European Union's General Data Protection Regulation (GDPR). It establishes guidelines for the governance and protection of personal data, including principles of transparency, accountability, and security (Greek Government Gazette, 2019).
- National Data Protection Authority of Greece (DPA), that provides guidelines and recommendations for data governance practices in line with the GDPR (Greek Government Gazette, 2006, 2019). It offers resources and templates to help organizations develop their data governance frameworks and ensure compliance with data protection regulations.

**Data Classification and Categorization**

Data classification and categorization refer to the process of organizing and labeling data based on predefined criteria, which may include data type, sensitivity, value, purpose, or any other relevant factors. The main objective is to systematically group and classify data to improve its management, security, and accessibility.

Data classification typically involves assigning different levels of classification or sensitivity to data. Commonly used classification levels include:

- Confidential/Highly sensitive data that requires the highest level of protection due to its sensitive nature, such as personally identifiable information (PII), financial data, or trade secrets.
- Internal/Restricted data that is intended for internal use within an organization and may have some restrictions on access or distribution.
- Public/Non-sensitive data, which is publicly available and does not contain any sensitive information.

Data Categorization on the other hand is based on various criteria, such as data type (text, images, audio, video, or structured/unstructured data), business function (finance, marketing, human resources, or operations) and regulatory compliance. These categorization criteria help in organizing data in a meaningful way, enabling efficient data management and retrieval.

Implementing data classification and categorization brings several benefits in data security, management, compliance and decision-making. By systematically organizing and labeling data, organizations can enhance their data governance practices and optimize the use and protection of their data assets.

Data Access and Security Policies
Data access and security policies are crucial for both the public sector and large organizations to ensure the protection, confidentiality, integrity, and availability of data, because they define the rules and procedures for accessing, handling, and safeguarding sensitive information. Some key considerations for data access and security policies may include:

- Access Control: Implementing strict access controls is essential to restrict data access to authorized personnel only. This can involve role-based access control (RBAC), where access rights are assigned based on job roles and responsibilities. It also includes the use of strong authentication mechanisms, such as passwords, multi-factor authentication, or biometrics (Best & Nelson, 2020).
- Data Encryption which is a vital security measure to protect data from unauthorized access, as it involves converting data into an unreadable format using encryption algorithms. Both data at rest (stored data) and data in transit (data being transmitted) should be encrypted to ensure its security.
- Data Handling and Storage and retention of data, which includes defining procedures for secure data storage, regular data backups, and secure data disposal when it is no longer needed.
- Data Monitoring and Auditing in regular basis that are essential to identify any suspicious or unauthorized access attempts and may involve monitoring log files, conducting periodic security audits, and implementing intrusion detection systems to detect and respond to security breaches promptly.
- Staff Training and Awareness that is crucial to provide training and awareness programs to employees regarding data security policies, best practices, and their responsibilities in safeguarding data.
• Incident Response and Data Breach Management that includes processes for reporting, investigating, mitigating, and recovering from data breaches, as well as communicating with stakeholders and regulatory authorities, if required.

Data Quality Assurance and Validation
Data quality assurance involves implementing processes, policies, and procedures to monitor and improve the quality of data. It aims to identify and resolve data issues, such as errors, inconsistencies, duplicates, and missing values. This helps in maintaining the overall data quality and ensuring that the data is fit for its intended purpose.

Data validation refers to the process of verifying the accuracy, validity, and compliance of data against predefined rules, standards, or requirements. It ensures that the data meets specific criteria and is suitable for its intended use. Data validation checks can include validation rules, format checks, range checks, and referential integrity checks to ensure data integrity.

Triple-Entry Accounting (TEA) has been proposed by Ian Grigg to increase transparency, accountability, and security in financial transactions. This method expands upon the traditional double-entry accounting system, which records transactions as debits and credits in two separate ledgers, by incorporating a third ledger as an independent verifier via a digitally signed receipt. The utilization of a digital signature provides evidentiary power to the receipt, thus reducing the accounting problem to one of the presence or absence of the receipt. The integrity issues associated with double-entry accounting can be addressed by allowing the parties involved in the transaction to share the records with an external auditor. This method can also be expanded to records stored on a Distributed Ledger Technology system (ie: Blockchain) while encryption mechanisms can be utilized to address the privacy issues (Sgantzos, Hemairy, Tzavaras & Stelios, 2023).

Techniques such as data profiling, data cleansing, and data validation checks are commonly employed to assess and enhance data quality. Data quality assurance and validation processes should be integrated into the overall data governance framework of an organization and involve collaboration among various stakeholders, including data owners, data stewards, and IT professionals. By incorporating data quality practices into the governance framework, organizations can ensure that data is managed effectively and consistently throughout its lifecycle.

The use of automated tools and technologies, such as data quality management systems, can streamline and enhance the data quality assurance and validation processes. These tools can help in identifying data issues, performing data cleansing tasks, and enforcing data validation
rules. They provide efficiency and accuracy in managing data quality, particularly in large and complex datasets. In Table 1, we can find a compilation of popular commercial big data validation tools. This table presents a comparison of these tools based on their operational environments, supported data sources, data validation capabilities, and examples of successful applications (Error! Reference source not found. Xie & Tao, 2016).

Table 2.1: A Comparative Analysis of Big Data Validation Tools and Key Participants

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation Environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Windows</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linux</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAC OSX</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debian</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CentOS</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ubuntu</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solaris</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VMware</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Java Environment</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intuitive user interface</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open-source ETL/Extract-Transform-Load</td>
<td>○</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>File Features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSV/TSV</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TXT Files</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Server Logs</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NoSQL</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twitter Firehole</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facebook Graph API/Thrift</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fitted Value Test</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFML</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JSON</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XML</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log4Log File</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIVE</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impala</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MySQL</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DB2</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PostgreSQL</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teradata</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sybase</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excel</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amazon Redshift</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data Retention and Disposal

Data retention policies are established to comply with legal and regulatory requirements, as well as to meet business needs. These policies define the duration for which data should be retained based on factors such as the type of data, its sensitivity, and the purpose for which it was collected (Tsesis, 2019). On the other hand, data disposal refers to the secure and permanent removal of data when it is no longer needed. This process ensures that data is properly destroyed to prevent unauthorized access or misuse. Data disposal methods may include physical destruction of storage media, such as shredding or degaussing, or secure deletion of digital data using specialized software.

Effective data retention and disposal practices are essential for maintaining data security, minimizing storage costs, and complying with privacy regulations. Organizations should establish clear policies and procedures for data retention and disposal, outlining the responsibilities of individuals or teams involved in managing data. Regular reviews and audits should be conducted to ensure compliance with the policies and to identify and address any data retention or disposal issues.

According to Data Protection Commission (n.d.), individuals have the right to request the erasure of their data, and controllers must respond to such requests without undue delay. However, there are exceptions, such as when processing the data is necessary for exercising the right of freedom of expression, compliance with a legal obligation, or the establishment, exercise, or defense of legal claims. The General Data Protection Regulation (GDPR) Articles 17 and 19 play a crucial role in safeguarding individuals' privacy rights, specifically focusing on the right to erasure.

Article 17, often referred to as the "right to be forgotten," empowers individuals to request the deletion or removal of their personal data when certain conditions are met. These conditions include situations where the data is no longer necessary for the purpose it was collected, the individual withdraws consent, or the data was processed unlawfully. Article 19 complements Article 17 by imposing an obligation on data controllers to communicate any erasure or rectification of personal data to recipients. This ensures that the right to erasure is not only exercised by removing the data from the controller's systems but also by extending the communication of such actions to others who may have received the data.

Methods

The research methodology employed in this study is a comprehensive and systematic approach aimed at investigating the intricacies of data management practices in public administration and large organizations. The methodology encompasses both qualitative and quantitative research techniques, providing a holistic understanding of the subject matter.
Data Collection
A thorough review of existing literature on data management practices, frameworks, and challenges in public administration was conducted. This involved an extensive exploration of academic journals, books, government reports, and reputable online sources. Several case studies were analyzed to glean insights into real-world implementations of data management strategies in public administration. These case studies were selected based on their relevance and the richness of information they provided.

Data Analysis
Qualitative data, including insights from literature reviews and case studies, underwent thematic analysis. This involved identifying key themes, patterns, and trends in data management practices. The accuracy and reliability of the study are contingent on the availability and accuracy of the data collected.

Conclusions and Suggestions
Effective data management practices are crucial for public administration and large organizations to ensure the proper handling of data throughout its lifecycle. These practices involve various activities such as data acquisition, storage, security, retrieval, dissemination, archiving, and disposal. Key takeaways from effective data management include the importance of data classification, categorization, and the utilization of tools and systems like database management systems (DBMS) and data governance frameworks.

To improve data management, it is recommended that public administration and large organizations establish comprehensive data management strategies and policies aligned with their goals and industry best practices. Investing in robust data management tools and technologies is essential for efficient data storage, access, security, and quality assurance. Additionally, clear data access and security policies should be developed, encompassing user authentication, authorization, and encryption measures.

Implementing data governance frameworks is crucial to ensure proper data classification, categorization, and the management of data throughout its lifecycle. Building a culture of data literacy and providing training to employees on data management practices, including data quality assurance and validation techniques, is also recommended. Despite the existence of data quality standards developed by ISO, the International Organization for Standardization (ISO/TS 8000-1:2011), it is still necessary for individual enterprises to establish their own specific standards and programs for ensuring the quality of their big data. This includes addressing the quality of both the big data itself and the application systems and services associated with it.

Wiig (2000) highlights the application of knowledge management in public administration, emphasizing the role of knowledge in shaping effective data management strategies. As
knowledge management intersects with data management, organizations should leverage both to enhance decision-making processes and foster innovation.

The Organization for Economic Cooperation and Development (OECD) employed an analytical framework to explore the various aspects of Data-Driven Public Sector (DDPS) that consists of three primary areas of focus. Firstly, there is a comprehensive model for data governance that encompasses multiple dimensions. Secondly, the application of data for public value is emphasized, highlighting the significance of utilizing data to deliver positive outcomes for the public. Lastly, the framework addresses the role of data in building public trust, recognizing the importance of maintaining trustworthiness and transparency in data practices. Within these three areas, there are 12 sub-dimensions that are proposed (Figure 3).

In the context of mismeasure in public administration, Lavertu (2016) highlights the challenges associated with "big data" and its potential impact on the accuracy and validity of public administration metrics. This underscores the importance of critically assessing data sources and methodologies to ensure the reliability of information used in decision-making processes.

Figure 3.1. 12 Dimensions of DDPS

Looking ahead, several future trends and areas for further research in data management have emerged. Exploring emerging technologies such as artificial intelligence (AI) and machine learning (ML) for data management tasks can bring advancements in data classification, data quality assessment, and automation of data management processes. The impact of data privacy regulations and ethical considerations on data management practices is another important area of study, as well as the use of blockchain technology for secure and transparent data management, along with the integration of big data analytics and data management practices. Finally, the potential of cloud computing and data storage technologies in enabling scalable and flexible data management solutions and the role of data management in supporting digital transformation initiatives in public administration are other significant areas for future exploration.
References