Analyzing Data Governance Components in Enterprise Information Systems: A Comparative Study

Ng Suang Joo
Faculty of Computing
Universiti Teknologi Malaysia
Skudai, Malaysia
ngjoo@graduate.utm.my

Ling Wan Yin
Faculty of Computing
Universiti Teknologi Malaysia
Skudai, Malaysia
lingyin@graduate.utm.my

***Abstract*—This study explores the data governance frameworks and highlights the requirement of particular methods for various organization types. The research emphasizes the value of simplified frameworks for small and medium-sized enterprises (SMEs) and how they engage in data governance discussions, drawing attention to the efficacy of the DAMA-DMBOK 2 framework in Human Capital Management. The study looks into the function of all-encompassing data governance frameworks in optimizing the value of business assets and promoting well-informed decision-making within the domains of enterprise information management and smart city infrastructure. It highlights the critical elements of successful implementation, which include organizational commitment, well-defined standards and procedures, and suitable technology.**

Keywords*—Data Governance, Data Quality, Data Management, Enterprise Information System, Data Governance Frameworks, SME, Smart City

I. Introduction

In the realm of a competitive industry, effective management of enterprise information systems (EIS) must be on the right track toward the success of an organization. At the core of an enterprise information system, there always exists the issue of data quality. Data governance has its impact in addressing these data quality issues. As suggested by DAMA International [6], data governance is the acquisition of authority and oversight over the handling of data resources, involving development, observation, and compliance. Therefore, data governance can be clarified as a structure containing policies and procedures for ensuring high-quality and secured data within an organization. The term 'Garbage In, Garbage Out' underpins the importance of high-quality source data in influencing decision-making, which causes businesses to be more inclined to establish a robust data governance framework within their organizations, regardless of any complexity and challenges that arise during the implementation process [5]. The need for data governance becomes even more critical given the significant growth in data volume, particularly in enterprise information systems. As organizations continue to gather enormous volumes of data involving structured, semi-structured, or unstructured data from various sources, the cost of ensuring high-quality data can be unexpectedly high. Organizations also will be exposed to the risk of data silos, inconsistencies, and inaccuracies inside their EIS if proper data governance does not occur in place.

Speaking of data governance, there are distinct variations in the components of a data governance framework. Those variations highlight the need for a comparative study to assess the effectiveness and how their best practices varied across different contexts. Thus, this paper aims to present the analysis results of the governance framework and provide insightful information about the requirements of the data governance framework within the context of EIS through a systematic review of relevant literature and case studies with the intention of this study to improve the data quality ease in the decision-making process.

II. Comparative analysis of previous studies

Prior research studies have examined various aspects of data governance frameworks and have discovered a significant variance in how they are implemented and designed. Considering this, Table 1 presents a comparative analysis of data governance components extracted from previous studies to focus on identifying the key components that might guide establishing strong data governance initiatives.

TABLE I. Comparison of data governance framework

Study	Research Area	Key Findings
I. F. Ruslan, M. F. Alby, and M. Lubis [4]	Human Capital Management (HCM)	Ten components to address data-related issues within HCM.
C. Begg and T. Caira [3]	Small and Medium Enterprises (SME)	Poorly implemented, frameworks might not be appropriate or scalable.
K. Paskaleva [2]	Smart City Infrastructure	Effective in changing culture surrounding data generation, collection, and utilization.
L. Cheong and V. Chang [5]	Enterprise Information Management	Comprehensive framework involving three component categories.

A. Human Capital Management (HCM)

Ruslan et al. [4] characterize Human Capital Management (HCM) as recognizing employees as valuable company assets or investments. The HCM integration with

Data Analytics (DA), also known as People Analytics (PA), is where the utilization of data to optimize HR initiatives. However, to address the rise of challenges related to data management practices, Peeters, Paauwe, and Van De Voorde introduced the People Analytics Effectiveness Wheel as displayed in Fig. 1, emphasizing governance's role in effective PA implementation, including Data Governance [8].

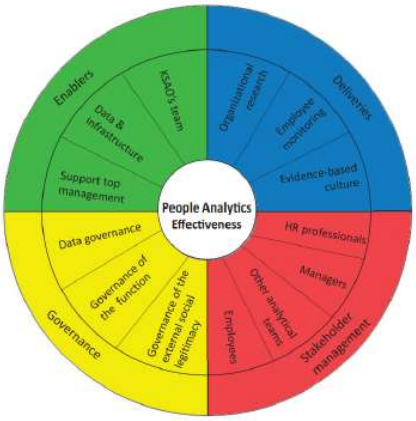

Fig. 1 People Analytics Effectiveness Wheel [4]

In line with this, the DAMA-DMBOK 2 framework is chosen as a structured strategy to address data-related issues within HCM. It encompasses 10 distinct data management perspectives–Data Architecture, Data Modelling & Design, Data Storage & Operations, Data Security, Data Integration & Interoperability, Document & Content Management, Reference & Master Data, Data Warehousing & Business Intelligence, Metadata, and Data Quality [4] as depicted in Fig. 2.

Fig. 2 DAMA-DMBOK 2 Wheel [4]

The adoption of each component of the framework focusing on Human Capital reveals a structured strategy to address data-related issues within the HCM domain as shown in Fig. 3.

Data Governance DAMA-DMBOK 2		
Data Management	**Focus in General**	**Focus in Human Capital**
Data Architecture	Policy on corporate IT Architecture masterplan in developing a data-driven ecosystem.	HC Data Policy will adopt from the main data architecture for developing HC Data Warehouse.
Data Modelling & Design	Policy on data modelling for serving the data on several functions.	The data is model based on the nature and condition of Human Capital Management.
Data Storing & Operational	Policy on how data will be stored, back up to and recovered.	Data storing will have its own mechanism based on how HC function Operates.
Data Security	Policy on how data will be secured from internal and external attacks.	The policy on data security must be expanded to the scope of HC Managers to their employees themselves.
Data Integration & Interoperability	Policy on how data will be integrated and transformed for consolidation in the data Hubs.	As HC Data grows following with the expanding size of the organization, it will need a Standard Procedure for data integration.
Document & Content Management	Policy on how contents is managed within documents. This applies on non-relational data.	Digitalized HC Documents (PDFs, Images, etc) should be managed carefully as it can be used for evidence during audit practices.
Reference & Master Data Management	Policy on how master data and reference data will be managed since it will impact on data integration and quality.	The scope of HC data authority for sharing practices and stewardship must be defined appropriately with HC business processes since it can affect data quality.
Data Warehousing & Business Intelligence	Policy on how Data Warehouse and Business Intelligence should be managed to support the organization's decision-making process.	Data driven HC practices will need a clear and concise Business Intelligence which is supported by developing Data Warehouse specialized for HC.
Metadata Management	Policy on how Metadata will be managed.	Metadata for HC arises from the nature of HC business processes which generates time dependent data.
Data Quality Management	Policy on how maintaining data quality.	Maintaining HC Data quality will need contribution from not only data stewards but also with HC data users since perceived quality differs on both sides.

Fig. 3 DAMA-DMBOK 2 Components in Human Capital [4]

B. Small and Medium Enterprises (SME)

According to the paper written by Carolyn Begg and Tom Caira, SME encounters several challenges regarding data governance. Firstly, while recognizing the benefits, SMEs often perceive the effort required for data governance as outweighing the benefits, partly due to challenges in gaining control over their IT systems and managing externally sourced data. Secondly, SMEs may not prioritize data governance as they fail to recognize the inherent value of their data or perceive it as being constrained by application packages, leading to IT systems dictating data governance practices. An illustration of this may be found in Fig. 4, where the enterprises explain how they comprehend the "five decision domains" that make up the Khatri & Brown framework. Thirdly, there is a lack of suitable data governance frameworks for SMEs, with existing frameworks often being complex and unsuitable for those with limited technical expertise and resources. Even the simplest frameworks, such as Khatri & Brown (2010), fall short when applied to SMEs. Lastly, SMEs may be overlooked by the data governance community, despite their significant economic contribution, potentially leading to regulatory challenges as data governance becomes increasingly important in the digital landscape. The authors advocate for greater attention to SMEs in data governance discussions and stress the importance of recognizing the value of data governance for their survival in the digital era, especially as regulatory requirements in this area are likely to increase across all sectors [3].

Decision Domain	Awareness and Understanding
Data Principles	All of the enterprises struggled to explain this term with most believing it to relate to external regulatory or compliance frameworks and only three of the enterprises providing responses that provided any focus on internal use and management of corporate data.
Data Quality	All of the enterprises identified this term as relating to the accuracy and integrity of their electronic data.
Metadata	None of the enterprises could offer any form of explanation for this term.
Data Access	Enterprises offered differing explanations for this term with some enterprises relating it to security and others explaining it in relation to the variety of ways in which a dataset can be accessed.
Data Lifecycle	Only one of the enterprises provided an explanation of this term that matched the definition provided in the framework

Fig. 4 Summary of responses to organizational awareness and understanding of decision domain terminology [3]

Fig. 5 provides an overview of the five decision domains and illustrative examples of the kinds of decisions that fall under each category made by Khatri & Brown.

Data Governance Domains	Domain Decisions	Potential Roles or Locus of Accountability
Data Principles • Clarifying the role of data as an asset	• What are the uses of data for the business? • What are the mechanisms for communicating business uses of data on an ongoing basis? • What are the desirable behaviors for employing data as assets? • How are opportunities for sharing and reuse of data identified? • How does the regulatory environment influence the business uses of data?	• Data owner/trustee • Data custodian • Data steward • Data producer/supplier • Data consumer • Enterprise Data Committee/ Council
Data Quality • Establishing the requirements of intended use of data	• What are the standards for data quality with respect to accuracy, timeliness, completeness and credibility? • What is the program for establishing and communicating data quality? • How will data quality as well as the associated program be evaluated?	• Data owner • Subject matter expert • Data quality manager • Data quality analyst
Metadata • Establishing the semantics or "content" of data so that it is interpretable by the users	• What is the program for documenting the semantics of data? • How will data be consistently defined and modeled so that it is interpretable? • What is the plan to keep different types of metadata up-to-date?	• Enterprise data architect • Enterprise data modeler • Data modeling engineer • Data architect • Enterprise Architecture Committee
Data Access • Specifying access requirements of data	• What is the business value of data? • How will risk assessment be conducted on an ongoing basis? • How will assessment results be integrated with the overall compliance monitoring efforts? • What are data access standards and procedures? • What is the program for periodic monitoring and audit for compliance? • How is security awareness and education disseminated? • What is the program for backup and recovery?	• Data owner • Data beneficiary • Chief information security officer • Data security officer • Technical security analyst • Enterprise Architecture Development Committee
Data Lifecycle • Determining the definition, production, retention and retirement of data	• How is data inventoried? • What is the program for data definition, production, retention, and retirement for different types of data? • How do the compliance issues related to legislation affect data retention and archiving?	• Enterprise data architect • Information chain manager

Fig. 5 Framework for data decision domains [7]

C. Smart City Infrastructure

Three important dimensions—data collection, use, and management—that have been previously recognized in research on data in smart cities are integrated in Fig. 6. Other dimensions that have been found include the project context, data generation, data identification, data sharing, and data heritage. These dimensions are arranged in a sequential arrangement into six clusters. The primary data governance decision domains that were determined through case studies during the first baseline stage were the focus of the Data Governance Survey. The survey looked at many facets of data and was carried out as part of the Triangulum project. It explored organizational, legislative, technical, and policy aspects that impact the creation and collection of data. It also looked at new methods for managing data, obstacles to data sharing within the project, and the line that should be drawn between management and governance in light of existing frameworks. Inquiries also touched on legacy challenges, the anticipated value of generated data to stakeholders, and the efficiency of teamwork in data generation and project management [2].

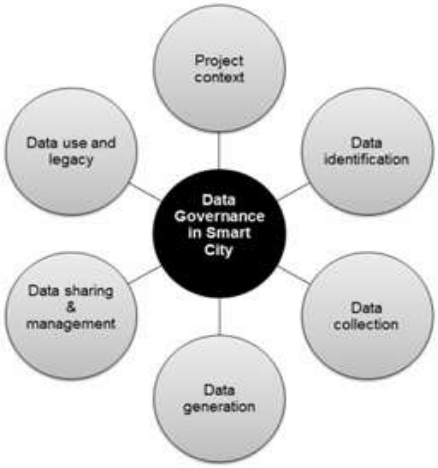

Fig. 6 Key pillars of data governance in smart city initiatives [2]

D. Enterprise Information Management

As mentioned in the case study, the initial data governance known as 'Governance via Stewardship', faces challenges within the data management team due to several reasons: (1) the absence of a mandate from senior executives results in a lack of authority to take action, (2) projects may not receive priority when senior executives fail to grasp their significance and (3) the absence of well-defined roles and responsibilities lead to ambiguity regarding the definitions of data stewards and data owners [5]. Therefore, the case study presents effective data governance known as the 'Governance via Governance' model, focusing on the three principal categories–Organisational Bodies and Policies, Standards and Processes, and Data Governance Technology as shown in Fig. 7.

Organisational Bodies and Policies	Standards and Processes	Data Governance Technology
- Governance Structure - Data Custodianship - User Group Charter - Decision Rights - Issue Escalation Process	- Data Definition and Standard (Meta data management) - Third Party Data Extract - Metrics Development and Monitoring - Data Profiling - Data Cleansing	- Metadata Repository - Data Profiling tool - Data Cleansing tool

Fig. 7 Data Governance Framework Components [5]

III. OPINION

The effectiveness of the DAMA-DMBOK 2 framework in Human Capital Management (HCM) may be significant or limited. Its structured framework includes an extensive range of data management that can be customized to meet the specific requirements of HCM such as data architecture, modeling, storage, security, integration, and quality management. Through those components focusing on Human Capital, an organization can protect their employee data security from breaches and reliability and guarantee data consistency and reliability, which facilitates the organization in decision-making with continuous up-to-date employee data across the organization's departments.

However, the designed framework for HC may be only suitable for application in larger, established organizations with structured data management practices where resource availability, organization maturity, and operational complexity are within their scope. Smaller and more agile

companies may find it challenging to adopt the entire comprehensive framework but may still profit from the individual DAMA-DMBOK 2 framework components with modifications of their unique requirements and limitations. In essence, organizational size, maturity, and adaptability determine the effectiveness of the DAMA-DMBOK 2 framework when handling data in any system. More developed data management companies can directly take advantage of the framework components.

Since SMEs frequently struggle with the complexity of IT control and believe that the effort required outweighs the advantages, they must address their data governance concerns. Frameworks that are simplified and adapted to the needs of SMEs may make it easier for them to put into practice efficient data governance procedures. Furthermore, SMEs must participate in data governance conversations to guarantee thorough rules and procedures.

The integration of multiple elements in the context of smart city data governance emphasizes the need for comprehensive approaches while also highlighting the complexity of urban data ecosystems. One particularly important tool in this effort is the Data Governance Survey from the Triangulum project, which provides insightful information about the potential and challenges associated with local data governance. Nonetheless, obstacles continue to exist, including legacy problems that make it difficult to seamlessly integrate new governance practices with current systems and the blurring of the boundaries between management and governance positions, which could make it difficult to define duties clearly. Furthermore, certain inefficiencies in collaboration, especially in the creation of data and project management, highlight the necessity of ongoing development and enhancement of cooperative methods. Despite these obstacles, encouraging sustainable and efficient municipal data governance requires teamwork and a greater comprehension of the complex structure of urban data ecosystems.

In the context of Enterprise Information Management (EIM), the comprehensive data governance framework presenting the three main categories of components is imperative in maximizing the business assets' value. Those component categories contribute to accountability and alignment with business initiatives. Organizational Bodies and Policies encompass the involvement of a Governance Structure developed based on the responsibilities of each member who participated in data governance, ensuring IT initiatives are strategically matched with business objectives.

Standards and processes assist organizations in guaranteeing data quality, mitigating risks, and optimizing the value derived from their data assets, thereby fostering business success and gaining competitive advantage with the help of clear guidelines and procedures. The data governance technology assists in better gaining insights into their data and thus introducing informed decision-making. Overall, these comprehensive components are necessary for supporting the continuity of EIM strategies where understanding and accountability of data are requirements to ensure effective data governance. The concepts and procedures that EIM employs are possibly transferable to other domains such as customer relationship management (CRM) with outstanding prospects. Nevertheless, the over-reliance on data governance technologies imposes the challenge of cost constraints, not to mention the investment for continuous maintenance and updates. Technology can sometimes be troublesome when it comes to situations where human intervention is manageable to do so. Therefore, while technology may be a catalyst in upholding data quality, organizations must always balance the use of technology and humans.

IV. Conclusion

When comparing the utilization of data governance components across different contexts of studies, it is evident that some components are tailored to meet industry-specific requirements and regulatory environments. However, the core principles of data governance remain relevant in all settings. For example, in Human Capital Management (HCM), data governance components must prioritize the security and confidentiality of employee data, while in a smart city, the focus may be on data sharing and citizen engagement. Regardless of the specific implementation and prioritization of data governance components, the ultimate objective is to ensure data security, integrity, and alignment with organizational goals. These objectives form the basis for an effective Enterprise Information System (EIS). Thus, organizations must carefully evaluate industry-specific characteristics and regulatory environments when designing and implementing their data governance frameworks to stay competitive and influential within the EIS domain.

References

[1] A. Alsousi and A. Shah, "DATA GOVERNANCE FOR SME: SYSTEMATIC LITERATURE REVIEW," *Journal of Information Systems and Digital Technologies*, vol. 4, no. 2, Dec. 2022, Available: https://journals.iium.edu.my/kict/index.php/jisdt/article/view/237

[2] K. Paskaleva, J. Evans, C. Martin, T. Linjordet, D. Yang, and A. Karvonen, "Data Governance in the Sustainable Smart City," *Informatics*, vol. 4, no. 4, p. 41, Nov. 2017, doi: https://doi.org/10.3390/informatics4040041.

[3] C. Begg and T. Caira, "Exploring the SME Quandary: Data Governance in Practise in the Small to Medium-Sized Enterprise Sector," *Electronic Journal of Information Systems Evaluation*, vol. 15, no. 1, pp. pp3-13–pp3-13, Jan. 2012, Available: https://academic-publishing.org/index.php/ejise/article/view/237/200

[4] I. F. Ruslan, M. F. Alby, and M. Lubis, "Applying Data Governance using DAMA-DMBOK 2 Framework: The Case for Human Capital Management Operations," *Proceedings of the 8th International Conference on Industrial and Business Engineering*, Sep. 2022, doi: https://doi.org/10.1145/3568834.3568866.

[5] L. Cheong and V. Chang, "The need for data governance: A case study," Jan. 2014, Accessed: Mar. 20, 2024. [Online]. Available: https://www.researchgate.net/publication/228966685_The_need_for_data_governance_A_case_study

[6] Dama International. "DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition)". Technics Publications, LLC, Denville, NJ, USA, pp. 71, 2017.

[7] V. Khatri and C. V. Brown, "Designing Data Governance," *Communications of the ACM*, vol. 53, no. 1, p. 148, Jan. 2010, doi: https://doi.org/10.1145/1629175.1629210.

[8] T. Peeters, J. Paauwe, and K. Van De Voorde, "People analytics effectiveness: developing a framework," *Journal of Organizational Effectiveness: People and Performance*, vol. 7, no. 2, pp. 203–219, Jun. 2020, doi: https://doi.org/10.1108/joepp-04-2020-0071.