MASTER DATA GOVERNANCE: IMPLEMENTING BEST PRACTICES THROUGH TECHNOLOGY AND TOOLS

Manykandaprebou Vaitinadin
Independent Researcher, Pondicherry, India
mkprebou1@gmail.com

Abstract

Master data governance, though not a new concept, has gained significant traction in recent years due to the increasing need for businesses to optimize costs, accelerate product development, ensure regulatory compliance, and maintain a competitive edge. Although master data governance has been around since the 1990s, the growing complexity of managing and maintaining high-quality data has become a major challenge for many organizations. Companies that effectively manage their data have a clear advantage, as data-driven decision-making leads to more efficient operations and better strategic outcomes. Conversely, poor data quality or mismatched information can hinder growth and derail business objectives. This paper explores the best practices for implementing master data governance, offering insights on how organizations can improve data accuracy, consistency, and reliability. By adopting these practices, companies can enhance their ability to make informed decisions, streamline operations, and drive business success with greater predictability and efficiency.

Keywords: Metadata, Master Data Governance, Master, Data Management, Data Challenges, Centralized Data, Framework, Best Practices MDM, Data Governance Tools.

I.   INTRODUCTION

Data governance refers to the formal process through which an organization manages its data assets, recognizing their critical role in driving business success. This process encompasses the establishment of processes, roles, and standards to ensure the efficient and effective use of data across the organization (1). The success of most organizations today is closely tied to the quality of their data management practices. When data is well-managed, it can significantly contribute to business growth, while poor data governance can impede progress and lead to inefficiencies. Effective data governance ensures the usability, integrity, availability, consistency, and security of data within an organization. Master data governance, a crucial subset of this process, focuses on the creation of policies for managing and maintaining critical master data. To address the challenges associated with centralized data management, several best practices are essential. These include data consolidation, the establishment of a master data governance team and operational model, mapping the flow of master data across production and consumption, governing shared metadata, and implementing policies to ensure data quality throughout the

entire data lifecycle (2). By adopting these practices, organizations can significantly enhance their data governance framework and unlock the full potential of their data assets.

Figure1: Master Data Management

II.   METHODOLOGY AND DESIGN

2.1 Challenges in Data Management

In many organizations, high-quality data is considered a strategic asset, providing a solid foundation for business growth. Quality data enables top management to make informed decisions, driving better outcomes for the business. However, critical data, particularly regarding customers, products, and partners, is often dispersed across various departments within the organization, creating significant challenges in managing this vital asset. This fragmentation leads to data being siloed, increasing the risk of misrepresentation, misinterpretation, and duplication, which ultimately hinders organizational performance (3).

To overcome these issues, organizations must centralize their data, consolidating it into a single, accessible point. By sharing data centrally, organizations can improve management efficiency and ensure that all authorized personnel can access the data when needed. This centralization also facilitates the identification of new patterns and insights, helping organizations explore new business opportunities, such as launching new product lines or services.

2.2 Centralized Data Management

A key best practice in master data governance is the consolidation of data across an organization. Data consolidation involves gathering and centralizing master data from various departments into a single repository. However, this process is not just about collecting data; it also includes identifying and cleaning duplicate or identical data entries. During consolidation, data objects are carefully scrutinized to ensure accuracy, and redundant copies are removed to

avoid unnecessary duplication in the central repository (4).

By consolidating data, organizations can significantly improve their business analytics capabilities. This centralized data provides a more accurate and comprehensive view, enabling the organization to deliver reliable business insights to stakeholders and enhancing decision-making across the enterprise.

2.3 Framework for Effective Big Data Governance

Big data governance involves the management and regulation of large data sets according to established rules and standards. As digital technology plays an increasingly central role in global development, particularly in the areas of social and economic progress, data collection and privacy have become critical concerns [5]. Effective big data governance frameworks provide structured methodologies for addressing these challenges and ensuring compliance.

Telecommunications companies, for example, have a significant opportunity to create new revenue streams through Big Data Analytics (BDA) solutions, driven by the surge in data traffic and shifts in customer behavior, particularly in light of the ongoing global health crisis [6]. The rise of digital data and ICT infrastructure has greatly enhanced the operational efficiency of both private and public organizations [7], enabling better decision-making and more informed strategies.

The Big Data governance framework incorporates various elements that are crucial to maintaining data quality. Unlike traditional governance models, this framework emphasizes the need for timely, reliable, relevant, and sufficient data services [8]. It also addresses a wide array of governance issues that require meticulous attention to ensure its effectiveness.

In an effort to evaluate existing frameworks, several studies on Big Data governance frameworks (e.g., BGF1 to BGF12) were compared against the ISO 8000 data governance standards. A comparison table (Table 2) outlines the alignment, where FR represents full representation and PR indicates partial representation. However, due to the unavailability of the BGF7 framework, a direct comparison could not be made.

Despite its promise, the rapid expansion of regional health information networks (RHINs) has faced setbacks, primarily due to the complexities associated with managing big data [9]. The practical implications of big data governance present both significant challenges and unique opportunities for organizations seeking to harness the power of large-scale data effectively [10].

2.4 Creating a Master Data Governance Team and Operating Model

To successfully manage master data governance, organizations should establish a cross-functional team that brings together individuals from various departments. Since business processes span multiple functions, it's essential for the team to have diverse expertise to effectively collaborate on creating and maintaining master data. A common pitfall in many organizations is entrusting the entire responsibility of master data governance solely to the IT department, which can lead to challenges. IT professionals may not fully understand the specific processes of other departments, making it difficult for them to address data issues in a

comprehensive manner (11).

To ensure a streamlined approach to managing master data, the governance team needs to adopt an operating model that includes clear policies for managing the data lifecycle, maintaining data quality, and enforcing rules for data management. In addition, a communication strategy should be developed to promptly notify relevant stakeholders about data errors or risks to data usability, allowing the organization to respond effectively and mitigate potential issues. This holistic approach can enhance the governance of master data across the organization.

2.5 Mapping the Creation and Use of Master Data

To ensure the effective use and quality of data across various departments, it is crucial to identify and map both the points of creation and the points of consumption for master data. The production points refer to the business processes responsible for generating data within the organization, while consumption points identify how and where the data is used by different functions. By documenting these processes, organizations can gain a clear understanding of how master data flows through the system, ensuring its availability and usability across multiple departments even when stored centrally.

Mapping the production and consumption of master data helps ensure that the data is continuously updated, maintained, and accessible. It also allows organizations to trace the data's journey from creation to consumption, ensuring that no data is overlooked or lost. Additionally, identifying these touchpoints enables the development of rules and guidelines for data governance, ensuring that master data is managed efficiently and remains accurate throughout its lifecycle. This comprehensive approach ensures that the organization's data remains reliable and supports decision-making across all functions (11).

2.6 Managing Shared Metadata

One of the key challenges in master data governance arises from inconsistencies and discrepancies in shared data, especially when it comes to duplication of data sets related to customers and products. To address this issue, it is essential to implement effective governance measures for shared metadata. Centralized management and oversight of this data are critical best practices to ensure its consistency and reliability.

The oversight team responsible for shared metadata governance should focus on managing key data elements, reference data, and associated rules to eliminate discrepancies. This governance process involves continuously monitoring the data to address challenges such as data similarity and normalization. By doing so, organizations can reduce the frequency of errors, improve data accuracy, and minimize the need for repetitive reconciliation efforts. This approach not only enhances the overall data quality but also strengthens the efficiency of data management strategies (11).

2.7 Developing Policies to Ensure Data Quality Throughout the Data Lifecycle

Maintaining high-quality data throughout its lifecycle is critical to effective data management.

Organizations must establish clear policies that govern every stage of the data lifecycle to ensure the integrity and quality of their data (12). These policies should cover data collection standards, governance procedures, and quality assessment protocols, aiming to eliminate inconsistencies, anomalies, and incomplete data.

By setting guidelines for proper data handling, these policies help reduce errors, prevent misuse, and ensure that data remains accurate and reliable at all times. Consistent enforcement of these policies fosters a robust data management framework that supports both operational efficiency and strategic decision-making.

2.8 Tools for Big Data Governance Apache Atlas

Apache Atlas is a key tool in managing big data governance, particularly within Hadoop ecosystems. It provides a comprehensive framework for metadata management, enabling the governance of Hadoop components through effective data governance practices. Apache Atlas utilizes the concept of data tagging to enhance security and ensure proper classification and management of sensitive information across big data environments (13).

III.   ALATION DATA CATALOG

The Alation data catalog is a centralized platform that consolidates data from multiple sources, providing a single access point for users to retrieve data. It facilitates data governance by managing activities such as customer engagement, updates, and collaboration within the data management process [14].
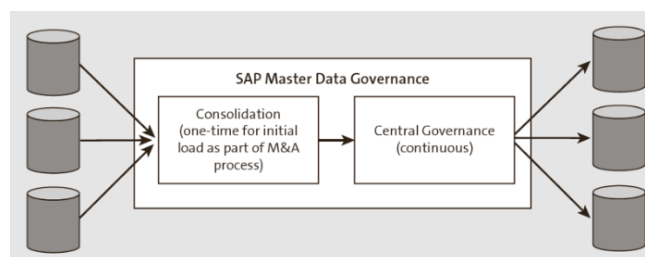

Figure 2: SAP Master Data Governance

3.1 SAP Master Data Governance

SAP Master Data Governance (MDG) helps organizations maintain high-quality data by implementing rules and regulations that govern data management. Text analysis highlights several key terms that are closely associated with Master Data Management (MDM) research, including: [15] master data, [16] data quality, [17] business intelligence, [18] business processes,[19] data integration,[21] big data, data governance, 8) information governance,[13] data management, and [14] product data [20]. Metadata plays a crucial role in identifying essential data within an organization, using various terminologies to facilitate easy access to relevant information.

IV. CONCLUSION

Organizations should prioritize a comprehensive approach to data management, recognizing that effective data management is key to their success. Master data governance must be an ongoing process rather than a one-off initiative,as both the business landscape and organizational data are constantly evolving. It is essential for organizations tofocus on maximizing Return on Investment (ROI) by involving all departments in master data management. By adhering to the best practices of data governance outlined, organizations can enhance their ability to thrive and stay competitive in today's fast-paced business environment.

REFERENCES

1. Oracle. "Enterprise Information Management: Best Practices in Data Governance." May 2011. http://egovstandards.gov.in/sites/default/files/oea-best-practices-data-gov-400760.pdf.
2. Loshin, D. "5 Best Practices for SAP® Master Data Governance." 2012. http://docs.media.bitpipe.com/io_10x/io_104906/item_536606/Winshuttle-5-Best-Practices-SAPMDG-whitepaper-EN.pdf.
3. ICF. "Data Governance Best Practices and Recommendations Report." May 2019. https://cdn.atlantaregional.org/wp-content/uploads/data-governance-best-practices.pdf
4. PAT Research. "12 Best Practices for Master Data Management." PAT RESEARCH: B2B Reviews, Buying Guides & Best Practices. Last modified September 20, 2021
5. M. Flyverbom, R. Deibert, and D. Matten, "The governance of digital technology, big data, and the internet: New roles and responsibilities for business," Business & Society, vol. 58, no. 1, pp. 3-19, 2019
6. M. Z. Kastouni and A. A. Lahcen, "Big data analytics in telecommunications: Governance, architecture and use cases," Journal of King Saud University-Computer and Information Sciences, 2020
7. A. Al-Badi, A. Tarhini, and A. I. Khan, "Exploring big data governance frameworks," Procedia computer science, vol. 141, pp. 271-277, 2018.
8. H. Y. Kim and J.-S. Cho, "Data governance framework for big data implementation with NPS Case Analysis in Korea," Journal of Business and Retail Management Research, vol. 12, no. 3, 2018.
9. Q. Li et al., "A framework for big data governance to advance RHINs: a case study of China," IEEE Access, vol. 7, pp. 50330-50338, 2019.
10. A. T. Ali Al-Badi, Asharul Islam Khan, "Exploring big data governance frameworks," Procedia Computer Science, pp. 271-277, 2018.
11. Valcheva, S. "20 Data Management Best Practices: Strategies That Work." Blog For Data-Driven Business. Last modified May 24, 2020. https://www.intellspot.com/data-management-best-practices/

12. Loshin, D. "Data Governance for Master Data Management and Beyond." 2013. https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/data-governance-for-MDM-andbeyond-105979.pdf.
13. D. Well, Introduction to Data Catalogs. Alation, 2019.
14. S. D. Neil D Lawrence, Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance. International Data Privacy Law, Volume 9, Issue 4, November 2019, Pages 236–252, 2019
15. S. Sagiroglu, D. Sinanc, "Big data: A review," in 2013 International Conference on Collaboration Technologies and Systems (CTS), 2013.
16. V. Morabito, "Big data governance," in Big data and analytics, 2015, pp. 83-104.
17. G. George, Haas, M. R., & Pentland, " Big data and management," Academy of Management JournalVol. 57, No. 2, pp. 64-72, 2014.
18. S. Soares, Big data governance. Information Asset, LLC. 2012.
19. A. T. Ali Al-Badi, Asharul Islam Khan, "Exploring big data governance frameworks," Procedia Computer Science, pp. 271-277, 2018
20. F. Haneem, R. Ali, N. Kama, and S. Basri, "Descriptive analysis and text analysis in systematic literature review: a review of master data management," in 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), 2017: IEEE, pp. 1-6
21. M. P. Atsushi Yamada, "Governance framework for enterprise analytics and data," in 2017 IEEE International Conference on Big Data, 2017.