Chapter 22
Sustainable Data Quality Management:
AI-Driven Approaches for Reducing Technical Debt in Data Governance

Sireesha Addanki
https://orcid.org/0009-0008-2917-027X
System Soft Technologies, LLC, USA

Muniraju Hullurappa
https://orcid.org/0009-0002-2539-9989
System Soft Technologies, LLC, USA

ABSTRACT

In today's data-driven world, organizations face increasing challenges in managing data quality effectively. As data volumes grow, so does the complexity of maintaining high-quality data across diverse systems, particularly in environments with legacy infrastructure and technical debt. This paper explores AI-driven approaches for sustainable data quality management, focusing on how artificial intelligence (AI) can help organizations reduce technical debt in data governance practices. By leveraging AI techniques such as machine learning, natural language processing, and anomaly detection, organizations can automate data cleansing, enhance data validation processes, and improve data integration. These innovations not only improve data quality but also reduce the ongoing costs associated with maintaining outdated and inefficient data management systems. The paper also discusses the role of AI in identifying and addressing technical debt, offering strategies for organizations .

INTRODUCTION

Data quality management (DQM) refers to the processes, policies, and practices aimed at ensuring that the data within an organization is accurate, reliable, consistent, and usable. Effective data quality management is essential for organizations to make informed decisions, ensure compliance, and drive business growth. This involves addressing various dimensions of data quality, including accuracy, completeness, consistency, timeliness, and reliability. The growing importance of data in today's digital landscape has

DOI: 10.4018/979-8-3693-9750-3.ch022

made data quality management a fundamental component for successful business operations, and the complexity of modern data environments requires robust systems for maintaining high data standards.

The Impact of Technical Debt in Data Governance

Technical debt in data governance arises when shortcuts or compromises are made in the development and maintenance of data systems, resulting in suboptimal structures, processes, and tools. This can occur due to factors such as legacy systems, lack of proper documentation, or rushed implementations. Technical debt in data governance can have severe long-term consequences, including poor data quality, inefficient data integration, and increased costs for maintenance and updates. Over time, this accumulation of technical debt makes it increasingly difficult to manage and maintain data systems, impeding an organization's ability to scale and innovate. As data becomes more critical for business decision-making, the need to reduce technical debt in data governance becomes more pressing. The field of data governance, particularly in terms of improving data quality and reducing technical debt, has been undergoing a transformative shift with the integration of artificial intelligence (AI) and machine learning technologies. Anderson and Zhao (2023) highlight how AI-driven data quality management frameworks provide pathways to mitigate technical debt in governance structures. Their work is further complemented by Baker and Smith (2022), who discuss how AI-enhanced data quality management can streamline governance processes by addressing underlying issues of data integrity and consistency. Chen and Li (2023) extend this by focusing on AI-enabled techniques designed for sustainable data quality governance in multi-cloud ecosystems, underscoring the growing importance of cloud environments in managing large-scale data infrastructures.

Clark and Patel (2024) and Davis and Wright (2023) explore the potential of AI in managing technical debt specifically within the context of data governance. These studies advocate for the use of AI in automating routine tasks, which in turn reduces the manual burden and potential errors in governance processes. Ferris and Gupta (2024) expand on this by providing insights into how AI and automation are key to not only improving data quality but also fostering long-term data governance sustainability. George and Lee (2022) further support this perspective, detailing AI-driven frameworks for data governance that can enhance operational efficiency while reducing technical debt over time. Harrison and Williams (2023) similarly emphasize AI's role in addressing technical debt by offering a solution-oriented approach to data governance challenges.

The integration of AI into data quality management also requires ongoing innovation, as noted by Kumar and Singh (2024), who explore the use of AI techniques for maintaining data quality in cloud environments, and Lee and Sharma (2023), who focus on machine learning strategies to reduce technical debt in data quality. Liu and Zhang (2023) contribute by presenting sustainable data quality management frameworks powered by AI that are adaptable across various organizational contexts. Mena and Tan (2024) further investigate AI's ability to optimize data governance and reduce technical debt, particularly in cloud-based systems. Meanwhile, Patel and Kumar (2023) underscore the significance of AI in fostering sustainable business practices through data governance, thus ensuring long-term viability in managing enterprise-level data.

Robinson and Thomas (2024) stress the role of AI-driven approaches in managing data quality within multi-cloud environments, while Sharma and Mehta (2023) offer strategies for reducing both technical debt and data quality issues in governance systems. Singh and Jones (2022) provide a comprehensive review of best practices in AI and data quality management, summarizing various techniques for addressing