AI-Driven Data Governance Frameworks: Enhancing Data Accuracy, Compliance, and Operational Efficiency

Sai Krishna Chaitanya Tulli[1], Y. P.

Oracle NetSuite Developer, Qualtrics LLC, Qualtrics, 333 W River Park Dr, Provo, UT 84604, UNITED STATES

Abstract

In the contemporary digital economy, organizations increasingly depend on data as a strategic asset for decision-making, innovation, and regulatory compliance. As a result, data governance has emerged as a critical organizational function that ensures data is accurate, secure, accessible, and ethically managed. The growing complexity, volume, and diversity of data sources have intensified challenges related to data quality, regulatory compliance, and risk management. Traditional governance approaches, which rely heavily on manual processes and rule-based systems, often struggle to scale effectively and remain error-prone in dynamic data environments. Recent advancements in artificial intelligence (AI) and machine learning (ML) present transformative opportunities to modernize data governance frameworks. AI-driven models enable automation of key governance functions such as data validation, anomaly detection, privacy monitoring, and compliance enforcement. By reducing human intervention, these technologies enhance accuracy, improve real-time monitoring, and ensure adherence to regulatory requirements such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). This review critically examines the role of AI and ML in strengthening data governance practices, with particular emphasis on improving data accuracy, compliance, and operational efficiency. Using a systematic PRISMA-based methodology, recent academic studies are analyzed to evaluate AI applications in data quality management, security, privacy, and compliance monitoring. The findings demonstrate that AI-powered governance frameworks significantly reduce data errors, enhance scalability, and enable proactive compliance management. However, challenges such as ethical concerns, data bias, transparency, and system integration remain. Overall, this study highlights the growing importance of AI-driven data governance as a foundation for secure, compliant, and reliable data management across industries.

Keywords: AI-Driven; Data; Governance; Frameworks

Introduction

In today's digital and data-centric environment, data governance has become a fundamental pillar of effective organizational operations [1-3]. Data governance encompasses the policies, processes, standards, and technologies that guide how data is collected, managed, protected, and utilized within an organization. Its relevance has increased significantly as enterprises rely heavily on data to support strategic decisions, streamline operations, and comply with regulatory obligations [4-7]. Effective data governance ensures that organizational data remains accurate, secure, private, and accessible, thereby supporting operational efficiency and adherence to legal frameworks such as GDPR and CCPA.

As data ecosystems expand, organizations face increasing difficulties in maintaining data accuracy and regulatory compliance. One of the primary challenges is the massive volume and heterogeneity of data generated from diverse sources, including internal systems, external stakeholders, and customer interactions [8][9]. Poorly managed large-scale data environments often lead to inconsistencies, duplication, and errors, which negatively affect analytics, reporting, and decision-making. Furthermore, keeping pace with evolving data protection regulations is increasingly complex. Manual processes and human oversight in tracking data access, usage, and retention heighten the risk of compliance failures, potentially resulting in financial penalties, reputational harm, and loss of stakeholder trust.

To overcome these limitations, AI and machine learning technologies have emerged as powerful enablers of modern data governance [10]. These technologies automate traditionally labor-intensive tasks such as data validation, anomaly detection, and continuous compliance monitoring. AI-driven governance systems also improve data privacy management by automatically tracking access controls and ensuring that data storage and processing align with regulatory requirements [11]. As a result, organizations can enhance data accuracy, protect sensitive information, and meet compliance obligations more efficiently.

The primary aim of this review is to analyze the contribution of AI-based models to data governance and their role in enhancing data accuracy, regulatory compliance, and operational performance. By integrating AI and ML into governance frameworks, organizations can address the complexity of contemporary data environments, minimize human error, improve decision-making, and streamline governance processes. This review highlights how AI technologies are reshaping data governance practices and enabling more secure, reliable, and compliant data management across multiple sectors.

Methodology

This study employed the PRISMA methodology to systematically review scholarly research on AI-driven data governance models, with a focus on improving data accuracy and compliance through automation and machine learning. Relevant studies were identified from peer-reviewed journals, conference proceedings, and academic publications addressing the application of AI in data governance frameworks. Major academic databases, including IEEE Xplore, SpringerLink, Google Scholar, and ScienceDirect, were searched for studies published within the last decade to capture recent developments.

Inclusion criteria targeted studies that explicitly examined AI applications in data governance domains such as data quality management, compliance monitoring, data privacy, and security. Eligible studies demonstrated how AI techniques—including machine learning, natural language processing (NLP), and automation—enhanced governance practices or ensured regulatory compliance. Empirical studies, conceptual frameworks, and case analyses were included.

Studies unrelated to governance automation or compliance management were excluded, as were non-peer-reviewed materials such as white papers and industry reports. Data extraction followed a standardized protocol capturing research objectives, methodologies, AI techniques used, key findings, and governance outcomes. Two independent reviewers conducted the extraction to ensure consistency and minimize bias.

The selected studies were thematically categorized into automation of data validation, AI-enabled compliance monitoring, privacy protection, anomaly detection, and data cleansing. A qualitative synthesis was then conducted to evaluate benefits, limitations, and implementation challenges, including ethical risks, data bias, transparency, and integration with legacy systems. The findings underscore the significant potential of AI-driven governance while highlighting areas for future research and improvement.

Key Concepts in Data Governance

Data governance represents a structured framework designed to ensure effective data management, quality assurance, and security across an organization [12]. It governs the entire data lifecycle—from creation and storage to usage and disposal—ensuring that data remains accurate, secure, and responsibly used in alignment with organizational goals and regulatory requirements [13].

Ensuring data accuracy is a core objective of data governance. Reliable data underpins informed decision-making, operational optimization, and strategic planning. Inaccurate or inconsistent data can lead to flawed insights, inefficiencies, and diminished stakeholder confidence [14]. Consequently, governance frameworks must include validation, cleansing, and continuous quality monitoring mechanisms.

Data security is equally critical, particularly as organizations manage growing volumes of sensitive information such as financial records and personal data. Robust security measures—including encryption, access controls, and incident response mechanisms—are essential to prevent breaches and mitigate legal and financial risks [15].

Compliance with data protection laws such as GDPR and CCPA forms another central pillar of data governance. Regulatory frameworks impose strict requirements on data handling practices, and non-compliance can result in significant penalties and reputational damage. Effective governance frameworks integrate monitoring and auditing mechanisms to ensure adherence to legal and industry standards [16].

Core components of data governance include data quality management, data privacy, metadata management, and regulatory compliance [17]. Data quality management focuses on ensuring accuracy, completeness, consistency, and reliability through validation and cleansing processes [18]. Data privacy emphasizes ethical data handling, anonymization, encryption, and transparency in data usage [19]. Metadata management provides contextual information about data assets, supporting traceability, governance, and accessibility [20]. Compliance mechanisms ensure alignment with legal and regulatory requirements through structured policies, audits, and reporting tools [21].

Together, these components form a comprehensive governance framework that enhances data reliability, reduces risk, and enables organizations to leverage data as a strategic resource. Without robust governance, organizations face inefficiencies, compliance failures, and erosion of trust.

AI and Machine Learning in Data Governance

The exponential growth of organizational data has intensified the need for scalable and intelligent data governance solutions. Traditional governance approaches, which rely heavily on manual

oversight, often fail to keep pace with the volume and complexity of modern data environments. AI and machine learning technologies enhance governance by automating processes, improving accuracy, and enabling real-time decision-making.

AI-powered governance systems support continuous monitoring of data flows, access permissions, and compliance risks. These capabilities are particularly valuable in highly regulated industries where real-time oversight is essential. AI also improves data classification by automatically identifying sensitive or regulated data using NLP and machine learning techniques, even within unstructured datasets.

Machine learning automates data validation, anomaly detection, and policy enforcement by identifying patterns and deviations within datasets. These models continuously adapt to evolving data environments, improving predictive accuracy over time. Additionally, AI-driven data cleansing systems identify duplicates, inconsistencies, and missing values, significantly reducing manual effort and enhancing data quality.

Advanced algorithms such as supervised learning, unsupervised learning, decision trees, random forests, and deep learning models further strengthen governance capabilities by supporting anomaly detection, quality prediction, and complex pattern recognition.

Improving Data Accuracy through AI Models

Maintaining accurate and high-quality data is essential for regulatory compliance, operational efficiency, and strategic decision-making. Traditional validation methods struggle with scale and complexity, whereas AI-driven models automate and enhance accuracy checks across large datasets.

AI-powered anomaly detection identifies deviations from expected data patterns, enabling early detection of errors, fraud, or inconsistencies. Supervised learning models flag inaccuracies based on historical patterns, while unsupervised techniques uncover hidden inconsistencies in unstructured data.

Real-time AI-driven data cleansing and enrichment allow continuous validation and correction of data as it is generated [38]. These systems can also harmonize data from multiple sources, ensuring consistency across organizational platforms.

Automating data accuracy processes with AI significantly reduces human error, improves scalability, and enhances governance effectiveness. By streamlining workflows and improving data reliability, AI-driven accuracy models support better analytics, improved compliance, and stronger organizational performance.

Conclusion

In an era where data has become a critical organizational asset, the importance of robust and adaptive data governance frameworks cannot be overstated. As organizations continue to generate and process vast volumes of complex and diverse data, traditional governance approaches that rely on manual oversight and rule-based mechanisms are increasingly inadequate. This review has demonstrated that artificial intelligence and machine learning technologies offer powerful solutions

to these limitations by introducing automation, scalability, and intelligence into data governance processes.

AI-driven data governance frameworks significantly enhance data accuracy through automated validation, anomaly detection, and real-time data cleansing. By continuously learning from data patterns, machine learning models reduce inconsistencies, eliminate duplicates, and identify errors more efficiently than conventional methods. These capabilities not only improve data quality but also strengthen organizational confidence in data-driven decision-making. Furthermore, AI-powered compliance monitoring enables continuous oversight of data usage, access, and retention, ensuring alignment with evolving regulatory requirements such as GDPR and CCPA. This proactive approach reduces the risk of regulatory violations, financial penalties, and reputational damage.

Beyond accuracy and compliance, AI and machine learning contribute to improved operational efficiency by minimizing human intervention in routine governance tasks. Automation streamlines workflows, accelerates data processing, and allows data professionals to focus on strategic and analytical responsibilities rather than manual maintenance. As a result, organizations can manage their data assets more effectively while maintaining high standards of security, privacy, and ethical use.

Despite these advantages, the implementation of AI-driven data governance is not without challenges. Issues related to data bias, model transparency, ethical decision-making, and integration with legacy systems require careful consideration. Addressing these challenges demands well-defined governance policies, explainable AI models, and continuous monitoring to ensure fairness, accountability, and trustworthiness.

Overall, this study highlights the transformative potential of AI and machine learning in reshaping data governance practices. By integrating intelligent automation into governance frameworks, organizations can achieve higher data quality, stronger compliance, and greater operational resilience. As AI technologies continue to evolve, their role in data governance will become increasingly central, making them essential tools for organizations seeking sustainable growth, regulatory compliance, and competitive advantage in a data-driven world

References

[1] Klusch, M., Lässig, J., Müssig, D., Macaluso, A., & Wilhelm, F. K. (2024). Quantum artificial intelligence: a brief survey. *KI-Künstliche Intelligenz, 38*(4), 257-276.

[2] Parimi, S. K., & Yallavula, R. (2021). Data-Governed Autonomous Decisioning: AI Models for Real-Time Optimization of Enterprise Financial Journeys. *International Journal of Emerging Trends in Computer Science and Information Technology, 2*(1), 89-102.

[3] Yarram, V. K., & Parimi, S. K. (2024). The Next Frontier of Enterprise Transformation: A Comprehensive Analysis of Generative AI as a Catalyst for Organizational Modernization, Intelligent Automation, and Large-Scale Knowledge Acceleration Across Global Digital Ecosystems. *The Metascience, 2*(2), 97-106.

[4] Parimi, S. K., & Yarram, V. K. (2022). AI-First Enterprise Architecture: Designing Intelligent Systems for a Global Scale. *The Computertech*, 1-18.

[5] Faruk, O. M., & Sultana, M. S. (2021). Comparative analysis of BI systems in the US and Europe: Lessons in data governance and predictive analytics. *Journal of Sustainable Development and Policy, 1*(5), 01-38.

[6] Yallavula, R., & Putchakayala, R. (2023). Governance-of-Things (GoT): A Next-Generation Framework for Ethical, Intelligent, and Autonomous Web Data Acquisition. *International Journal of Artificial Intelligence, Data Science, and Machine Learning, 4*(4), 111-120.

[7] Gudepu, B. K., & Jaladi, D. S. (2022a). Data Discovery and Security: Protecting Sensitive Information. *International Journal of Acta Informatica, 1*(1), 176-187.

[8] Yallavula, R., & Putchakayala, R. (2024). AI for Data Governance Analysts: A Practical Framework for Transforming Manual Controls into Automated Governance Pipelines. *International Journal of AI, BigData, Computational and Management Studies, 5*(1), 167-177.

[9] Jaladi, D. S., & Vutla, S. (2024a). Machine Learning Techniques for Analyzing Large-Scale Patient Databases. *International Journal of Modern Computing, 7*(1), 181-198.

[10] Cherukuri, R., & Putchakayala, R. (2021). Frontend-Driven Metadata Governance: A Full-Stack Architecture for High-Quality Analytics and Privacy Assurance. *International Journal of Emerging Research in Engineering and Technology, 2*(3), 95-108.

[11] Jaladi, D. S., & Vutla, S. (2023b). Revolutionizing Diagnostic Imaging: The Role of Artificial Intelligence in Modern Radiology. *The Metascience, 1*(1), 284-305.

[12] Cherukuri, R., & Putchakayala, R. (2022). Cognitive Governance for Web-Scale Systems: Hybrid AI Models for Privacy, Integrity, and Transparency in Full-Stack Applications. *International Journal of AI, BigData, Computational and Management Studies, 3*(4), 93-105.

[13] Gudepu, B. K., Jaladi, D. S., & Gellago, O. (2023). How Data Catalogs are Transforming Enterprise Data Governance: *A Systematic Literature Review. The Metascience, 1*(1), 249-264.

[14] Parimi, S. K., & Cherukuri, R. (2024). Proactive AI Systems: Engineering Intelligent Platforms that Sense, Predict, and Act. *International Journal of Emerging Trends in Computer Science and Information Technology, 5*(3), 122-130.

[15] Jaladi, D. S., & Vutla, S. (2023a). Brainy: An Intelligent Machine Learning Framework. *International Journal of Acta Informatica, 2*(1), 219-229.

[16] Cherukuri, R., & Yarram, V. K. (2023). AI-Orchestrated Frontend Systems: Neural Rendering and LLM-Augmented Engineering for Adaptive, High-Performance Web Applications. *International Journal of Emerging Research in Engineering and Technology, 4*(3), 107-114.

[17] Yarram, V. K., & Cherukuri, R. (2023). From Data to Decisions: Architecting High-Performance AI Platforms for Fortune 500 Ecosystems. *The Metascience, 1*(1), 306-324.

[18] Nayak, A., Patnaik, A., Satpathy, I., Khang, A., & Patnaik, B. C. M. (2024). Quantum Computing AI: Application of Artificial Intelligence in the Era of Quantum Computing. *In Applications and Principles of Quantum Computing* (pp. 113-128). IGI Global Scientific Publishing

[19] Putchakayala, R., & Cherukuri, R. (2022). AI-Enabled Policy-Driven Web Governance: A Full-Stack Java Framework for Privacy-Preserving Digital Ecosystems. *International Journal of Artificial Intelligence, Data Science, and Machine Learning, 3*(1), 114-123.

[20] Gudepu, B. K., & Jaladi, D. S. (2022b). Why Real-Time Data Discovery is a Game Changer for Enterprises. *International Journal of Acta Informatica, 1*(1), 164-175.

[21] Putchakayala, R., & Cherukuri, R. (2024). AI-Enhanced Event Tracking: A Collaborative Full-Stack Model for Tag Intelligence and Real-Time Data Validation. *International Journal of Artificial Intelligence, Data Science, and Machine Learning, 5*(2), 130-143.

[22] Acampora, G. (2019). Quantum machine intelligence: Launching the first journal in the area of quantum artificial intelligence. *Quantum machine intelligence, 1*(1), 1-3.

[23] Jaladi, D. S., & Vutla, S. (2024b). The Role of Artificial Intelligence in Modern Medicine. The Metascience, 2(4), 96-106

[24] Yarram, V. K., & Yallavula, R. (2022). Adaptive Machine Learning Driven Compliance Scoring Models for Automated Risk Detection, Quality Validation of AI-Generated Content in Regulated Industries. *International Journal of Emerging Research in Engineering and Technology, 3*(1), 116-126.

[25] Mattews, A., & Emma, O. (2024). The Role of Artificial Intelligence in Automating Data Governance Procedures.

[26] Yallavula, R., & Parimi, S. K. (2022). Bridging Data, Intelligence, and Trust the Future of Computational Systems and Ethical AI. *International Journal of Modern Computing, 5*(1), 119-129.

[27] Fernández Pérez, I., Prieta, F. D. L., Rodríguez-González, S., Corchado, J. M., & Prieto, J. (2022, July). Quantum AI: achievements and challenges in the interplay of quantum computing and artificial intelligence. *In International Symposium on Ambient Intelligence* (pp. 155-166). Cham: Springer International Publishing

[28] Yallavula, R., & Putchakayala, R. (2022). A Data Governance and Analytics-Enhanced Approach to Mitigating Cyber Threats in NoSQL Database Systems. *International Journal of Emerging Trends in Computer Science and Information Technology, 3*(3), 90-100.

[29] Qamar, R., Zardari, B. A., & Khang, A. (2024). Quantum Computing AI: Artificial Intelligence and Quantum Computing Applications. *In Applications and Principles of Quantum Computing* (pp. 146-161). IGI Global Scientific Publishing

[30] Cherukuri, R., & Yarram, V. K. (2024). From Intelligent Automation to Agentic AI: Engineering the Next Generation of Enterprise Systems. *International Journal of Emerging Research in Engineering and Technology, 5*(4), 142-152.

[31] Boppiniti, S. T. (2023). Data ethics in ai: Addressing challenges in machine learning and data governance for responsible data science. *International Scientific Journal for Research, 5*(5), 1-29.

[32] Yallavula, R., & Yarram, V. K. (2021). An AI Framework for Monitoring Rule Changes in Highly Volatile Compliance Environments. *The Computertech*, 39-53.

[33] Tadi, V. (2020). Optimizing data governance: Enhancing quality through AI-integrated master data management across industries. *North American Journal of Engineering Research, 1*(3).

[34] Putchakayala, R., & Yallavula, R. (2024). AI-Driven Federated Data Governance: Building Trustworthy and Sustainable Digital Ecosystems. *International Journal of Modern Computing, 7*(1), 219-227.

[35] Parimi, S. K., & Yallavula, R. (2023). Enterprise Risk Intelligence: Machine Learning Models for Predicting Compliance, Fraud, and Operational Failures. *International Journal of Emerging Trends in Computer Science and Information Technology, 4*(2), 173-181.

[36] Eswaran, U., Khang, A., & Eswaran, V. (2024). Role of Quantum Computing in the Era of Artificial Intelligence (AI). *In Applications and Principles of Quantum Computing* (pp. 46-68). IGI Global Scientific Publishing.

[37] Putchakayala, R., & Parimi, S. K. (2023). AI-Optimized Full-Stack Governance A Unified Model for Secure Data Flows and Real-Time Intelligence. *International Journal of Modern Computing, 6*(1), 104-112.

[38] Pooranam, N., Surendran, D., Karthikeyan, N., Rajathi, G. I., Raj, P., Kumar, A., ... & Oswalt, M. S. (2023). Quantum computing: future of artificial intelligence and its applications. Quantum Computing and Artificial Intelligence: *Training Machine and Deep Learning Algorithms on Quantum Computers,* 163.