Journal of Artificial Intelligence Practice (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/jaip.2025.080201
ISSN 2371-8412 Vol. 8 Num. 2

From Data Governance to Data Intelligence Governance: Transforming Enterprise-Level Data Asset Management

Bowen Ma

Tencent Technology (Shenzhen) Co. Ltd, Shenzhen, Guangdong, 518000, China

Keywords: Digital era, integrating AI, efficiency, analytics

Abstract: In the digital era, traditional data governance frameworks centered on rule - enforcement are struggling to meet the dynamic demands of modern enterprises, which restricts the efficiency of data utilization. This paper introduces a novel "data intelligence governance" framework. By integrating AI technologies with systematic governance practices, it creates a symbiotic relationship between governance and intelligent analytics. Empirical evidence shows remarkable improvements in data accessibility, analysis efficiency, and decision - making agility. Case studies demonstrate efficiency gains of up to 40%. This approach paves the way for scalable, self - service data asset management, enabling enterprises to transform into agile, data - driven organizations.

1. Introduction

The exponential growth of data has made robust data governance strategies essential for ensuring data quality, security, and compliance. However, conventional data governance models, with a narrow focus on rule - enforcement, are unable to adapt to the evolving business landscape characterized by dynamic market trends and emerging technologies. Recent advancements in large language models (LLMs) and automation technologies offer new opportunities to shift from passive governance to proactive "data intelligence governance". This paper explores this paradigm shift through Tencent's Cloud BI (ChatBI) and Ola data governance platform, demonstrating how the integration of governance standards with intelligent automation can create a virtuous cycle that enhances both data quality and analytical productivity.

Theoretical Foundation: The Resource-Based View (RBV)

The Resource - Based View (RBV) theory posits that a firm's sustainable competitive advantage stems from its unique resources and capabilities that are valuable, rare, inimitable, and non - substitutable. High - quality data, when effectively managed and transformed into actionable intelligence, fits this description. For example, the insights generated through ChatBI and Ola provide a competitive edge by enabling data - driven decision - making, highlighting the strategic importance of data in the digital age.[1]

2. Literature Review

2.1 Data Governance: From Compliance to Value Creation

Data governance involves establishing policies, roles, and processes to manage the entire data lifecycle. Traditional governance models primarily prioritize data trustworthiness and regulatory compliance, often through rigid rule - enforcement. For instance, in the early days of data governance in the financial industry, the focus was mainly on meeting regulatory requirements such as Basel III for banks, ensuring that data on capital adequacy and risk exposure was accurate and compliant.

However, modern paradigms are moving towards governance - as - a - platform. In this model, standardized processes and unified metadata enable self - service analytics. Frameworks like DataOps advocate for collaborative, agile workflows to align IT and business teams. A study by Gartner showed that organizations implementing DataOps principles experienced a 30% increase in data - driven decision - making speed within two years. Despite these advancements, challenges remain. For example, understanding the context of data in a multi - cultural and multi - domain enterprise setting is still difficult. A survey by McKinsey found that 70% of companies struggle with integrating domain - specific knowledge into their data governance frameworks.

The Theory of Constraints (TOC) in Data Governance

The Theory of Constraints (TOC) identifies bottlenecks in processes and advocates for their alleviation to improve overall efficiency. In data governance, bottlenecks such as slow data - quality checks can be a major issue. For example, in a large e - commerce company, manual data - quality checks were taking days, delaying critical business decisions. The Ola platform addressed this by automating data - quality checks. By using machine - learning algorithms to detect anomalies in data, it reduced the data - quality check time from days to hours. Research has shown that applying TOC in data governance can lead to a 25 - 40% improvement in data - processing efficiency.[2]

2.2 Intelligent Analytics: Augmenting Human Expertise

Large language models (LLMs) have revolutionized business intelligence (BI) tools. They enable natural language processing (NLP) and automated insight generation. Systems like ChatBI can reduce query latency by up to 70% through techniques such as retrieval - augmented generation (RAG) and prompt engineering. For example, a marketing team can simply ask "What is the sales trend of our new product in the last quarter?" in natural language, and ChatBI can quickly provide relevant data and analysis.

Data intelligence focuses on extracting actionable insights to support decision - making, leveraging tools like data catalogs, predictive analytics, and machine learning. However, challenges persist. Achieving accurate contextual understanding is still a problem. In a study by MIT Sloan Management Review, 80% of surveyed companies reported difficulties in getting context - relevant insights from their data intelligence tools.

Bayesian Inference in Data Intelligence

Bayesian inference enhances the reliability of data - driven insights by updating probabilities based on new data. In ChatBI, Bayesian methods incorporate prior knowledge and real - time interactions to make more accurate predictions. For example, in an e - commerce personalized recommendation system, Bayesian inference combines historical user purchase behavior with real - time browsing data. A research paper in the Journal of Machine Learning Research found that e - commerce platforms using Bayesian inference in their recommendation systems saw a 15 - 25% increase in user engagement.

3. Methodology: A Hybrid Governance - Intelligence Framework

Our proposed framework integrates two complementary systems:

1) Governance Backbone: Inspired by Tencent Ola's DataOps principles, emphasizing standardized modeling, unified metrics, and closed - loop governance.

2) Intelligence Engine: Leveraging ChatBI's LLM - based architecture for natural language querying, automated SQL generation, and insight synthesis.

3.1 Governance Design Principles

Standardization: Defining assetization criteria, such as metadata completeness and lineage transparency, aligns data production with governance goals. Complete metadata and transparent lineage ensure data integrity and accountability. For example, in a media company, standardized metadata for video content (including title, genre, production date, etc.) made it easier to manage and analyze large volumes of video data.

Automation: Employing Data Fabric architectures creates a unified data mesh, enabling dynamic discovery and access. This simplifies data integration and management, improving efficiency. A global enterprise with offices in multiple countries found that after implementing a Data Fabric - based data governance solution, the time spent on data integration across different departments was reduced by 50%.

Feedback Loops: Implementing governance engines, such as Ola's Quality Center, monitors compliance and triggers corrective actions, maintaining data quality and reliability. In a healthcare organization, the Quality Center continuously monitored patient data quality. When it detected inconsistent patient age entries, it automatically triggered an alert and provided suggestions for correction.

3.2 Intelligence Enablement

Intent Understanding: Addressing ambiguous user queries through intent optimization and RAG-enhanced LLMs reduces misinterpretation risks. For example, if a user asks "Show me the best-selling products", the system can clarify whether the user means best - selling in terms of quantity, revenue, or profit.

Task Automation: Deploying Text2SQL generators and visualization recommenders reduces manual coding efforts, making data more accessible. A small-to-medium-sized enterprise's business users were able to access and analyze data without relying on IT staff after implementing these tools.

Continuous Learning: Refining models through user feedback and domain - specific knowledge injection improves performance and context awareness. A financial services firm injected industry-specific knowledge about new financial regulations into its data intelligence model, which improved the accuracy of risk assessment.

4. Case Studies

4.1 Tencent Cloud BI: Democratizing Data Analysis

ChatBI addresses key pain points in traditional BI:

Natural Language Interaction: Resolves 85% of ambiguous queries through intent-clarification modules, improving user experience. For example, a salesperson asks "How's our performance?" ChatBI can ask for more details like "In which region?" or "For which product line?" to provide accurate answers.

Domain Knowledge Integration: Uses RAG to incorporate business glossaries, reducing hallucination risks by 60%. In a manufacturing company, ChatBI can understand and analyze terms like "production yield" and "defect rate" accurately.

Efficiency Gains: Automates SQL generation and visualization, cutting report - creation time from hours to minutes. A market research team used to spend 3 - 4 hours creating a sales performance report. With ChatBI, they can now generate the same report in less than 15 minutes.

Application Details: ChatBI has an intuitive user interface that is accessible to non - technical users. It also supports multi - language queries, which is beneficial for global enterprises.

Challenges and Coping Strategies: One challenge was handling extremely complex business logics. The development team addressed this by collaborating with domain experts to train the model with more complex business scenarios. Another challenge was data security when integrating with multiple data sources. They implemented strict access controls and encryption mechanisms.

4.2 Tencent Ola: Building a Governed Data Lakehouse

Ola’s governance framework ensures data reliability through:

Unified Metrics Store: Standardizes 90% of core metrics, eliminating duplication and discrepancies. In a telecommunications company, Ola unified metrics related to call quality, such as call drop rate and signal strength, across different departments.

Automated Lineage Tracking: Maps data flows across 500 + tables, enabling impact analysis in less than 5 seconds. When a data source for customer billing data was updated, Ola could quickly show which billing reports and downstream analytics would be affected.

Cost Optimization: Identifies 20% redundant storage through anomaly detection algorithms. A large-scale cloud storage service provider used Ola to clean up redundant data, saving significant storage costs.

Application Details: Ola provides a comprehensive dashboard for data governance administrators to monitor data quality, lineage, and compliance in real - time.

Challenges and Coping Strategies: During the implementation, data migration from legacy systems was a major challenge. The team developed a custom data migration tool to ensure data integrity. Another challenge was resistance from some employees to change their data - handling processes. The company organized training sessions and change - management programs to address this.

5. Innovation Pathways

5.1 Synergizing Governance and Intelligence

Governance - Driven Intelligence: Using policy - compliant data catalogs to train LLMs ensures regulatory adherence. For example, in the pharmaceutical industry, data catalogs that comply with FDA regulations are used to train LLMs for drug development data analysis.

Intelligence - Augmented Governance: Deploying AI for automated compliance checks and anomaly detection reduces manual audits by 50%. A multinational corporation used AI - powered anomaly detection to identify potential financial fraud in its expense reports.

5.2 Scalable Architectures

Modular Toolchains: Combining DataOps pipelines with LLM agents enables end - to - end automation. A software development company integrated its CI/CD pipeline with LLM - based code analysis tools for automated code review.

Hybrid Cloud Deployments: Leveraging Kubernetes and serverless architectures ensures elastic scalability. An online gaming company used hybrid cloud deployments to handle sudden spikes in user traffic during game launches.

5.3 Future Directions

Explainable AI: Enhancing trust through transparent reasoning chains. One way to achieve this is by developing visual explanation tools. For example, in a credit - scoring system, a visual graph can show how the AI model arrived at a particular credit score for a borrower, displaying the factors considered and their weights. Potential challenges include the complexity of developing such tools for complex deep - learning models and the need for domain - specific knowledge to interpret the explanations.

Edge Intelligence: Enabling real - time governance at the IoT edge using lightweight LLMs. To implement this, techniques for model compression and optimization are needed to run LLMs on resource - constrained edge devices. For example, in a smart city project, edge devices on traffic lights can use lightweight LLMs to analyze real - time traffic data and adjust light timings. Potential challenges include ensuring data security and privacy at the edge, as well as dealing with limited network connectivity.

6. Conclusion

The integration of data governance and intelligent analytics represents a paradigm shift towards self-sustaining data ecosystems. Tencent's ChatBI and Ola platforms demonstrate that governance standards and AI capabilities are mutually reinforcing. By adopting this hybrid approach, enterprises can achieve:

Faster Time - to - Insight: Automated pipelines reduce analysis cycles by 65 - 80%.

Higher ROI: Governed data assets yield 30 - 50% cost savings through optimized storage and reuse.

Regulatory Agility: Dynamic compliance frameworks adapt to evolving legal requirements.

As AI continues to mature, embedding governance into intelligent systems will lead to the creation of autonomous, data - driven organizations, driving innovation and competitiveness in the digital age.

References

[1] Barney, J. B. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99 - 120.
[2] Goldratt, E. M. (1984). The goal: A process of ongoing improvement. North River Press.