Infosys topaz

VIEW POINT

FROM DATA CHAOS TO AI-DRIVEN INSIGHTS: THE PATH TO ENTERPRISE SUCCESS

Abstract

Data has become one of the most valuable assets for businesses worldwide. Yet, despite massive investments in cloud infrastructure and artificial intelligence (AI) driven analytics, many enterprises struggle to organize, manage, and derive actionable insights from their data. Companies continue to face challenges related to data silos, fragmented governance frameworks, and ineffective deployment as well as utilization of AI.

Successful data strategies entail treating data as a key business asset, streamlining data movement, and shifting to an industry-driven open framework for interoperability. Open standards, AI-driven automation, and robust governance frameworks are key factors in making data accessible, trustworthy, and actionable. As a result, AI, data governance, and the cloud have become critical pillars of modern business operations.

This paper explores the challenges enterprises face in organizing the data and the role of AI in data utilization. It highlights the characteristics of successful data-driven organizations and the growing importance of governance and compliance. It further describes how organizations can transition from merely managing data to extracting real business value from it.

databricks

Infosys®
Navigate your next

Introduction

For several years now, enterprises have been struggling to harness the full potential of their data. While some organizations have managed to leverage artificial intelligence (AI) for valuable, actionable insights from data, most are mired in fragmented systems operating out of silos. Industry reports suggest that only 15% of organizations have successfully leveraged data to drive innovation, gain a competitive advantage, and fuel business transformation[1]. These companies prioritize interoperability, AI readiness, and robust governance[2].

Others, however, are stuck in a cycle of data duplication, compliance struggles, and outdated infrastructure, preventing them from achieving meaningful AI-driven insights.

As AI – particularly generative AI (GenAI) and agentic AI – continues to evolve, companies must rethink their approach to data management. Organizations that fail to establish a strong data foundation might struggle to reap the benefits of AI in the future, while those that embrace open standards, automation, and a streamlined architecture will gain a significant advantage.

The State of Data Management: Challenges and Opportunities

Enterprises often face significant challenges in implementing effective data management strategies, with data fragmentation remaining one of the most persistent obstacles. Data is often scattered across different systems, departments, and storage locations. This makes it difficult to access, analyze, and secure data. Compliance and regulatory pressures also create additional complexities. Laws such as the General Data Protection Regulation (GDPR) and the Digital Operational Resilience Act (DORA) require organizations to adopt standardized governance frameworks. However, fragmented data environments make it difficult to implement these frameworks effectively across multiple jurisdictions.

Many organizations also prioritize data engineering over deriving insights, leading to an imbalance in investment. Companies often invest heavily in building data pipelines but fail to focus on extracting meaningful business value. This results in time being frittered away on infrastructure and data plumbing rather than making their data trustworthy, accessible, and actionable. Most enterprises are also ill-prepared for the shift toward multimodal data—an integration of diverse data types such as text, images, audio, and structured data. For instance, in retail, AI models may need to analyze customer reviews (text), product photos (images), and in-store foot traffic data (sensor feeds) together to generate accurate insights into buying behavior and inventory planning. This capability is critical for AI applications.

What Sets the Successful 13% Apart?

The few organizations that excel in data utilization treat data as a core business function. They embed data-driven decision making at every level, enabling faster and more strategic operations. They adopt advanced AI ML techniques, integrating GenAI and agentic AI to accelerate automation and innovation.

A key characteristic of this success is simplification. Rather than being burdened by redundant tools, these organizations streamline their technology architecture while maintaining advanced capabilities. This efficiency allows them to focus on innovation rather than infrastructure maintenance.

Unlike their unsuccessful counterparts, these companies are multicloud-ready and prioritize interoperability, ensuring seamless data movement across platforms. They optimize their technology stacks by reducing redundancies while maintaining efficiency. Furthermore, they foster a strong culture of data literacy, empowering employees to actively engage with data to drive business value. As a result, these organizations consistently outperform their peers.

As they advance their data strategies, they also take proactive steps to ensure their data practices are aligned with regulatory requirements. They recognize the importance of strong data governance to support their innovation and AI-driven efforts.

Data Governance and Compliance

Regulatory requirements are constantly evolving, making data governance more important than ever. While GDPR has been in place for several years, many companies continue to struggle with fragmented compliance strategies. The introduction of DORA in January 2024 has added another layer of complexity, pushing organizations to further refine their governance frameworks in order to ensure improved compliance.

One of the biggest challenges is the fragmented nature of regulations. With different rules across 37 European countries and varying state-level regulations in the United States, multinational enterprises must navigate a complex regulatory environment. While traditional data governance, like GDPR, has established clear compliance measures, AI governance remains ambiguous, with no clear regulatory fines for non-compliance.

To address these challenges, successful enterprises are establishing common governance frameworks, observability measures, and control mechanisms, favoring transparency and explainability. Standardized compliance strategies help organizations maintain regulatory alignment without stifling innovation.

Unlocking the Potential of Data in Large Organizations

Successful organizations often leverage a data lakehouse architecture and 74% of global chief information officers report having a lakehouse in their estate3. It is a model built on open standards and principles that are designed to systematically eliminate data silos and enable interoperability. Open standards play a crucial role by offering flexibility and reducing reliance on proprietary systems, which often create inefficiencies.

Standardizing data formats enhances data lineage and discoverability, making it easier to track the origin and transformation of datasets. Additionally, open frameworks simplify multi-jurisdictional policy enforcement. This helps in allowing businesses to apply governance rules dynamically across different markets. Importantly, these rules are increasingly being expressed in open formats, such as UC Metrics. This creates greater transparency to guide adherence as well as enable comparison and alignment across different levels, whether national, regional, industry-specific, or ecosystem-wide. Although transitioning to this model can be complex, organizations that embrace open-source principles, AI-driven automation, and interoperability achieve better business outcomes.

The Role of AI and GenAI in Data Utilization

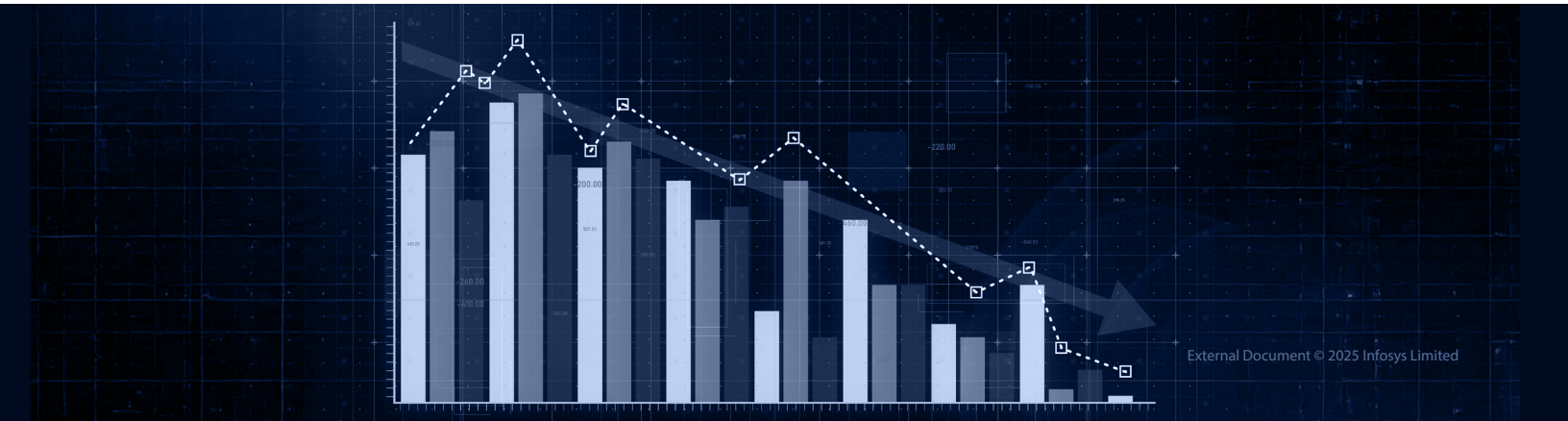
AI is transforming how enterprises manage data, but its effectiveness depends on a strong data foundation. Acting as a productivity catalyst, AI automates complex tasks such as code generation, document analysis, and compliance checks. At Infosys, for example, we have achieved an 80% increase in coding efficiency through AI automation.

AI also acts as a tool for valuable insight generation, swiftly identifying hidden patterns in vast datasets. However, organizations must tread with caution while making AI-driven decisions due to persisting concerns over trustworthiness and reliability. To help employees, customers, and stakeholders develop greater confidence in AI's capabilities, organizations must prioritize AI transparency, robust governance, and validation mechanisms to ensure credibility.

The Physical Nature of Data

Despite its digital nature, data has significant real-world implications. Just as transferring a large file between devices can take time, organizations frequently struggle with data movement across suppliers, partners, and customers. Latency, governance requirements, and operational inefficiencies are bottlenecks that slow this process down.

Every additional copy of data introduces delays, requiring further verification to maintain integrity. Legacy systems amplify this inefficiency, often leading to multiple copies and endless checks that slow down business operations, particularly in older organizations. It can sometimes take as long as six months to calculate the cost of a product under development, due to proprietary data locked in inefficient legacy systems. By adopting open-source technology and standardized formats, businesses can reduce unnecessary duplication and improve efficiency while maintaining governance and security. Optimizing their data architecture allows organizations to streamline operations and unlock faster, more effective data-driven decision-making.



Democratizing AI: Risks and Guardrails

Open standards are central to democratizing AI and data. While democratizing AI and data through open source models fosters innovation, it also introduces risks. It is hence critical to balance accessibility with security. Open-source data storage offers flexibility, but organizations must ensure sensitive content remains protected through robust, enterprise-level governance frameworks.

Further, AI governance must evolve beyond traditional document policies to include fine-grained control over AI training datasets. Standardized governance will ensure that AI remains transparent and compliant with evolving regulations across industries and geographies. Interoperability is a key success factor that can enable AI systems to function seamlessly within global enterprises, regardless of geographic or operational boundaries.

Conclusion

Data is no longer just an IT concern; it is the foundation of modern enterprises. However, most organizations continue to grapple with fragmented data, compliance challenges, and inefficient AI utilization. Companies that prioritize data literacy, adopt open standards, and streamline their architectures will gain a competitive edge. AI and GenAI present significant opportunities for businesses to enhance efficiency. But without a solid data foundation, these technologies cannot deliver meaningful insights. The future of data management lies in democratization through interoperability, where organizations leverage industry-driven open frameworks, robust governance mechanisms, and AI-driven automation to unlock real value from their data. Enterprises that embrace these principles will drive innovation and outperform peers to achieve greater business impact and growth.

References

1. Databricks | Building a high-performance data and AI organization
2. Forbes | Understanding Enterprise AI Readiness: It's All About Your Data
3. Databricks | Databricks named a Leader in the 2024 Forrester Wave for Data Lakehouses

Authors

Dael Williamson, EMEA CTO, Field Advisory and Engineering, Databricks

Rajan Padmanabhan, CTO, Data Analytics and AI, Infosys.

For more information, contact askus@infosys.com

Infosys®
Navigate your next