



eBook

How Data Leaders Empower Teams through Governed Data and AI

Unlocking AI-driven governance for success

**Where data
& AI come to** 



Contents

Introduction	3
AI Is Propelling Data Management's Evolution	3
Data Under Scrutiny: The Tough Interrogation by AI	4
The Rise of Modern Data and AI Governance	5
The 3 Pillars of Modern Data Governance	6
- 1. Managing Risk and Complying with Regulatory Mandates	6
- 2. Enabling Trusted Data Sharing and Democratization	8
- 3. Facilitating Intelligent Data Quality and Observability	10
Adopt a Holistic, Adaptive and Intelligent Approach to Deliver Trusted Data	11
Conclusion	13
About Informatica	14

Introduction

AI is fast becoming one of the most powerful drivers for accelerating digital business. The transformative impact of AI is compelling companies to reinvent ways of working with data across all parts of their value chain including customer experience, supply chain, marketing and sales, human resources and more.

However, these opportunities come with inherent risks. AI systems and applications need to be de-risked from bias, protected from data privacy and data security vulnerabilities, and must conform to legal and ethical considerations. Therefore, before organizations fully commit to their **AI adoption** initiatives, it is critical to take a pause and thoroughly examine the cornerstone of AI – data.

AI Is Propelling Data Management's Evolution

AI adoption is pushing organizations to evolve their data management maturity. In fact, data management is cited as a top technological inhibitor to AI/ML deployment.¹ Adoption demands emphasis on scalability, quality and governance of data available for AI systems. At the same time, renewed attention must be given to securing and efficiently accessing data and AI pipelines.

¹ <https://cioinfluence.com/machine-learning/new-study-reveals-data-management-is-a-top-challenge-in-the-ai-revolution/>

² https://forms.workday.com/en-us/reports/digital-acceleration-considered-approach-to-change/form.html?step=step1_default

To support AI applications with **trusted data**, data security, data privacy and **data governance** are increasingly important for chief information officers (CIOs) and IT leaders.

Research suggests that 52% of IT leaders highlight compliance, protection and/or privacy as their top IT investment areas.²

IT teams that rely on legacy technologies, fragmented tooling and manual data management processes will find it increasingly difficult to fuel AI initiatives with the reliable data that they need. However, AI-enabled data management platforms can help IT teams modernize their **data management** capabilities to support AI and decentralize the power of reliable AI-driven decision making across the enterprise.

In this eBook, we will explore the contemporary challenges of data that IT teams need to address and how CIO and IT teams can implement modern data and AI governance to accelerate successful business outcomes.

Data Under Scrutiny: The Tough Interrogation by AI

As promising as AI technology adoption does sound, organizations are realizing that for better results with AI, they need better data inputs. AI is focusing today's organizational readiness to manage and govern data at scale and speed under a spotlight by asking IT teams more difficult questions than they are historically equipped to handle. For example, these include:

- How do we ensure compliance with ever-changing AI and ethical data use regulations?
- Do we understand all the risks associated with exposing business-critical and sensitive data, and how do we safeguard it as it gets democratized?
- How can we maximize efficiencies in data operations to support scaling out our AI initiatives?
- How can we achieve visibility into the health of our data pipelines for AI to grow our digital business by removing barriers to adoption?

AI applications require large volumes of diverse data for training. But with large volumes of data comes larger responsibility. IT teams struggle to feed AI applications with appropriate and relevant data. The reason is simply because they are unable to govern such large volumes, variety, and velocity of data in accordance with regulatory and company policy mandates.

Not only do IT teams need to make large-scale data available for AI applications, but they also need to ensure that the data is reliable and accurate. The inability to continuously monitor the health of data across data pipelines, and to identify and rectify data anomalies early, can result in AI applications delivering inaccurate and unreliable outcomes that erode trust. Furthermore, to put AI to work effectively, IT teams also need to optimize the availability and accessibility of data by removing data dependencies without compromising the integrity of data.

To fuel AI systems with high-quality data at a scale and speed that AI can thrive on, traditional data governance is not enough. CIO and IT teams need to assess their company's data readiness for embracing AI and re-calibrate their ability to provide **governed data management** for AI success.

The Rise of Modern Data and AI Governance

Today, meeting the demands of a data-driven enterprise requires managing risk and compliance against an ever-changing regulatory landscape. Additionally, enabling trusted data sharing within a growing data community and achieving observability into the accuracy and explainability of data have become essential aspects of this complex landscape.

IT teams need to be enabled to monitor how data is generated, processed and consumed for analytics and AI while it remains trustworthy and compliant with policies. This requires a modern data and AI governance approach that involves three essential pillars:

1. Managing risk and complying with regulatory mandates
2. Enabling trusted data sharing and democratization
3. Facilitating Intelligent data quality and observability

These pillars help ensure data availability, reliability and protection to provide the foundation of trusted data and models needed for success with AI initiatives. Now, let us explore how CIOs can operationalize these three pillars of **modern data governance**.



Risk & Compliance

Manage risk and comply with regulatory mandates and policies



Analytics & Data Sharing

Open data for business with actionable insights and automation



Intelligent Data Observability

Optimize availability, performance, capacity – cost-effectively and efficiently

The 3 Pillars of Modern Data Governance

1. Managing Risk and Complying with Regulatory Mandates

CIOs operate at the intersection of data and how it is processed and used across the organization. As a result, they are responsible for establishing data policies and empowering teams with the foundational capabilities for regulatory compliance and data security. CIOs need to enable IT teams with complete visibility and control over how data is obtained, processed, and protected. This process must happen not only within the organization, but also externally, to be able to deploy safeguards that ensure adherence to privacy laws and security mandates.

Increasing volumes of data, complexities of a data landscape, increased risk from security vulnerabilities and ever-changing and new regulations make it challenging to protect data with legacy technologies. Adoption of enterprise AI has created additional risk considerations that include providing model transparency, enabling sensitive data use and mitigating biases in data.

Unsurprisingly, only one in five (21%) of IT leaders report having above 75% governance over their data collection systems and applications.³

Data-driven companies must adopt modern **data governance** that helps address the need to protect data and reduce potential risks without slowing down innovation. Advanced AI-enabled data governance solutions offer automated data discovery and classification as well as automatic enforcement of relevant data policies and procedures for appropriate use that help maintain policy compliance and limit risk exposure.

³ <https://www.wvntv.com/business/press-releases/ein-presswire/622606282/survey-finds-enterprise-cios-cisos-are-concerned-about-data-privacy-security-compliance-in-data-collection-process/>

The 3 Pillars of Modern Data Governance (Continued)

Critical capabilities required for managing risk and compliance management:

Data discovery. You cannot protect the data that exists outside of your awareness. With data scattered across diverse on-premises and cloud environments, locating and correctly identifying data presents a formidable challenge. Solutions that can provide complete visibility and transparency prevent data slipping through the cracks undetected and unprotected.

Data catalog. Understanding your data is equally important. A **data catalog** serves as a centralized inventory of all assets across an organization. By helping to classify data at scale and providing metadata insights, modern data catalogs can enforce applicable policy compliance that minimize the chances of data misuse. Additionally, a data catalog helps track the data use and documents transformations to supplement compliance efforts for timely regulatory reporting.

Data lineage. A clear view of data flows is critical towards ensuring data is protected and compliant as it propagates from source systems across applications. Automated **data lineage** enables IT teams to monitor changes at every stage of a data asset's lifecycle, helps streamline compliance processes and proactively identifies potential data issues and anomalies that may arise.

Automated governance workflows. Operationalizing governance frameworks with automation enables companies to scale the enforcement of consistent policies across enterprise applications by eliminating manual, repetitive tasks. It also empowers IT teams to adopt continuous monitoring by consistently flagging compliance risks.



The 3 Pillars of Modern Data Governance (Continued)

2. Enabling Trusted Data Sharing and Democratization

According to a survey by Bloor Research, 67% of IT leaders responded that access to useful data was restricted because of data governance mandates.⁴ The same study also highlighted that 60% of the respondents were unable to use more than 40% of the data.⁵

To get maximum value out of data and AI investments, companies must foster a data-driven culture. This entails ensuring that all data stakeholders can easily access trusted data for decision-making when they want and in the way they want.

CIOs play a pivotal role in promoting this culture of allowing data assets to be easily discoverable, trustworthy and accessible. In their role as the technology executive for the company, they need to collaborate with business leaders, such as chief data officers, chief information officers, chief marketing officers and chief human resources officers, to empower their teams with the trusted data and necessary access required for uncovering data-driven business insights.

At the same time, CIOs also must find a balancing act between keeping data open for business and risking the integrity and security of data. This is where reliable and confident **data sharing** and **data access management** capabilities are fundamental to a modern data governance approach.

Intelligent, integrated and automated methods to democratize data safely can boost data availability rapidly by combining data governance workflows, policy orchestration and data access management, while delivering a frictionless data access experience for data consumers.

⁴ <https://www.datanami.com/this-just-in/privitar-majority-of-data-leaders-unable-to-use-more-than-40-of-data-due-to-security-and-privacy-concerns/>

⁵ <https://www.datanami.com/this-just-in/privitar-majority-of-data-leaders-unable-to-use-more-than-40-of-data-due-to-security-and-privacy-concerns/>

The 3 Pillars of Modern Data Governance (Continued)

Critical capabilities for enabling data sharing:

Data marketplace. A governed **data marketplace** acts as a business-friendly storefront for data consumers to discover curated data assets from diverse locations, fully understand their context and usability, as well as access them in alignment with data sharing policies in a self-service manner.

Data access management. Ensuring the right people have access to the right data for the right purpose is crucial for IT teams before they can confidently loosen their grip on data and AI for wider organizational use. Data access management provides necessary controls. By automatically applying security and privacy protection across the data landscape, IT teams can unleash data while data remains compliant with policies for appropriate access and use.

Collaboration: With embedded collaboration tools, typically found within a data marketplace, data producers and consumers can collaborate in real time without depending on a mediator that is often from IT. This approach is crucial in establishing shared responsibility when defining the quality of relevant data. It also provides the structure needed to harness tribal knowledge. Embedded collaboration tools help to improve data literacy as data products are reviewed, recommended and prioritized for optimal business outcomes.

Automated data delivery. Data assets need to be delivered timely, safely and efficiently. With automated data delivery, data engineering teams can streamline delivery processes, pre-define default options for trusted data products, and help ensure data is delivered in the format required with optimal quality and appropriate policies enforced.



The 3 Pillars of Modern Data Governance (Continued)

3. Facilitating Intelligent Data Quality and Observability

Only when supported with high-quality data can AI be leveraged effectively to unlock new insights, drive efficiencies and foster innovation, with reliability and trust in business outcomes.

With the volume, diversity and number of data sources only expected to grow exponentially, **data quality** challenges are going to be compounded when attempting to scale out data pipelines. Data needs to be monitored for quality not only at the source, but across data pipelines during transformation. Further, with AI now in the picture, the onus to deliver high-quality data for training, updating and operating AI models at scale is even higher when considering the risks and rewards during adoption.

This can only be achieved by bringing data quality and data governance into the same conversation. CIOs need to empower their teams with integrated data quality tools. These tools should apply automation effectively to identify data quality issues and anomalies early. Additionally, observability tools must measure and remediate policy compliance and performance-level expectations to maintain data health.

Critical capabilities for facilitating intelligent data quality and observability:

Data observability. This capability provides the necessary tools and insights to effectively manage data operations and mitigate risks, breaking down barriers to maintaining healthy data pipelines. It ensures that data remains reliable, accurate and accessible in a timely fashion. **Data observability** can enhance data governance practices and help optimize data pipelines and data workflows by providing insights into performance metrics, data latency and resource utilization.

Data profiling. Before data can be integrated and used into any CRM, ERP or business application, its contents, quality and structure need to be validated. Data profiling allows you to assess data sources for completeness and accuracy. It can also help proactively identify data issues before they get used for erroneous business insights and decisions.

Data quality automation. Meeting the data quality requirements of AI initiatives by using manual processes to cleanse data is not feasible. Modern data governance solutions lean on the power of AI/ML to analyze large scale data sets, detect irregularities and provide insights for remediation activities. Operating at scale demands automation to be ingrained into the **heart of any modern solution**.

Adopt a Holistic, Adaptive and Intelligent Approach to Deliver Trusted Data

As digital businesses scale, the way in which data gets handled is becoming fundamentally altered. Support for modern business and technology initiatives that include real-time data analytics, cloud-based solutions, embedded AI, decentralized architectures and many more are pushing data leaders and their teams to reimagine how they manage and govern data to grow their business successfully.

In a relatively short time, the role of data governance has gone from being primarily focused on meeting regulatory compliance mandates to today becoming a central function of an organization's data-first journey by ensuring data availability, reliability and trust to fuel a next-generation data strategy.

CIOs can use this as an opportunity to transform their own role as a facilitator of data-first culture. Modern data governance solutions can help CIOs realign their data strategies to business objectives, plugging technology gaps.

These solutions also prepare the organization to make the most of their data by establishing a data foundation that is adaptive, holistic and intelligent.

Adaptive. Static control-oriented, hard-wired, one-size-fits-all data governance is neither sustainable nor desirable. Traditional data governance prescribes rigid and passive approaches that inhibit innovation. When organizations are faced with a rapidly changing technology environment, volatile regulatory requirements and evolving dynamics within the organization, modern data governance solutions provide IT teams the flexibility and adaptability required to evolve with the needs of the organization. This is achieved through an AI-enabled, automated approach that accelerates an ability to scale.

Intelligent. Even if an organization's data governance approach is comprehensive but takes enormous effort and time to operationalize, it is expected to be a significant roadblock to value. As the amount of data, diversity of data types and sources and potential use cases of data across the enterprise grow, data governance becomes progressively difficult without the help of AI and automation. By adopting AI-powered modern data governance solutions, organizations can save time and effort by simplifying tedious tasks such as data cleansing, data preparation, and democratizing and sharing data.

Adopt a Holistic, Adaptive and Intelligent Approach to Deliver Trusted Data (Continued)

Holistic. Organizations can no longer look at data management disciplines in silos. Modern data governance helps build a broad yet simple **metadata management**-based foundation to operationalize a host of capabilities. These include data discovery, data quality, data privacy, data sharing, and more, that are seamlessly integrated to solve current and future use cases. Modern data governance at scale demands both ease of adoption for these capabilities and the ability to expand as data strategy evolves.

Consider a simple example of a hospital administrator in a healthcare organization who needs to access patient health records to schedule follow-up visits. This may require data quality to verify the accuracy of a data set such as address information; data security to ensure compliance with regulations such as HIPPA; privacy controls and anonymization to safeguard electronically protected health information (ePHI) and other PII; authorization to access partial data aligned to role and context; a delivery mechanism to make the data available, along with policy enforcement including terms of use. Note that even this simple request entails multiple capabilities working together at the back end to deliver trusted data.

Conclusion

Catering to diverse data needs at the speed of business is always a daunting task. With added pressures of institutionalizing and operationalizing advanced technologies such as AI, and delivering ROI, this demand becomes even more overwhelming without the right support.

CIOs are increasingly involved in setting the vision and direction for data-driven business transformation. This positions them at the ideal vantage point to assess and plug gaps in the holy trinity of strategy, technology and culture with regards to effective data management.

To support a data-driven enterprise effectively, a fragmented data governance approach will prove inadequate. IT teams need to be empowered with modern data governance solutions that combine risk management, data quality and observability, and data sharing capabilities. This approach can maximize the value of data across the entire value chain for healthy and scalable data pipelines that fuel AI, analytics, customer experience and similar modern business initiatives.

Equipped with AI-powered data governance solutions, IT teams can build a better data foundation with the right capabilities that in turn can empower data consumers to use data with confidence and facilitate a data-driven culture.

About Us

Informatica (NYSE: INFA) brings data and AI to life by empowering businesses to realize the transformative power of their most critical assets. When properly unlocked, data becomes a living and trusted resource that is democratized across your organization, turning chaos into clarity. Through the Informatica Intelligent Data Management Cloud™, companies are breathing life into their data to drive bigger ideas, create improved processes, and reduce costs. Powered by CLAIRE®, our AI engine, it's the only cloud dedicated to managing data of any type, pattern, complexity, or workload across any location — all on a single platform.

IN19-4667-1223

© Copyright Informatica LLC 2023. Informatica and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and other countries. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners. The information in this documentation is subject to change without notice and provided "AS IS" without warranty of any kind, express or implied.

[informatica.com](https://www.informatica.com)

Where data & AI come to



Worldwide Headquarters
2100 Seaport Blvd,
Redwood City, CA 94063, USA
Phone: 650.385.5000
Fax: 650.385.5500
Toll-free in the US: 1.800.653.3871

[informatica.com](https://www.informatica.com)
[linkedin.com/company/informatica](https://www.linkedin.com/company/informatica)
twitter.com/Informatica

CONTACT US