

# Promote Responsible AI with AI-Powered Data Management

## Bridge the Trust Deficit with Responsible AI

As enterprises begin to unleash unprecedented value from AI, they can encounter growing risks. These include hallucinations, where incorrect information is presented as fact, biases from systemic distortions in statistical results, exposure of privacy and security, such as intellectual property infringement, algorithmic inaccuracies and insufficient explainability.

Without a responsible AI framework, organizations face increased liability, including reputational damage, loss of customer trust, financial losses and regulatory penalties. According to McKinsey's research, 44% of organizations indicated they have already experienced at least one negative consequence, such as data inaccuracy, privacy exposure or security.<sup>1</sup> Further, 62% of line-of-business users stated that they trust their organization's data only somewhat for decision-making.<sup>2</sup>

Organizations need to integrate and process large-scale, diverse datasets for training AI models, which helps manage the quality and integrity of data when governing an increasingly complex data and AI landscape. They must enact robust compliance measures to stay ahead of emerging AI regulations such as the **EU AI Act**. These all pose significant challenges for companies to adopt AI confidently.

## Key Benefits

- Build trust among stakeholders to accelerate AI adoption
- Improve the reliability of AI outcomes and help prevent hallucinations
- Reduce compliance risks and create a safer environment for AI innovation
- Help ensure data and AI operations are explainable and accountable
- Promote secure collaboration to help prevent data and AI misuse

<sup>1</sup> <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>

<sup>2</sup> <https://research.esg-global.com/reportaction/515201716/Marketing>

## Responsible AI Needs Integrated, AI-Powered Data Management

Informatica's **CDO Insights 2024: Charting a Course to AI Readiness** suggests that 58% of companies will need five or more data management tools to support generative AI (GenAI) challenges.<sup>3</sup> Fragmented tooling and niche data management solutions make it harder to get data ready for AI at the speed and scale required to drive reliable AI outcomes.

Companies need scalable, integrated **data management** solutions to handle increasingly vast and complex datasets. Solutions that leverage AI to help ensure data consistency, data quality and compliance empower organizations to foster transparency, **build trust** with stakeholders and enhance the reliability of AI outcomes while reducing risks.

## Enable Reliable, Trustworthy and Responsible AI with Informatica

Informatica's AI-powered, modern cloud data management helps create a trusted data foundation to drive responsible outcomes from GenAI applications. By enabling the availability and accessibility of high-quality, safe and protected data, the Informatica **Intelligent Data Management Cloud™** (IDMC) helps ensure trusted AI outcomes while mitigating privacy, security and compliance risks introduced through AI.

IDMC enables the scalability and flexibility needed to adapt to AI requirements without compromising performance. It provides comprehensive services to operationalize responsible AI initiatives with trusted data. Capabilities across data cataloging, data integration, data quality, master data management, data governance, data sharing, access management and data privacy help ensure that data in AI models is used responsibly and in alignment with organizational values.

### These capabilities include:

- Comprehensive data discovery and integration services to find and deliver the diverse data sets necessary for training AI models while minimizing bias.
- Data quality management and observability to empower companies to manage data health and help ensure AI models are trained on high-quality data.
- Metadata-driven data catalog and data governance capabilities to enable data intelligence with transparency and understanding and help enforce policy compliance.
- A governed data marketplace to democratize validated, curated data assets, offering better transparency of the data supply chain.
- Data access management to implement fine-grain controls and help prevent inappropriate access or misuse of data.
- Access management, data masking and anonymization to help ensure privacy is maintained while using sensitive data for AI.

Master data management to help create a single source of truth of business-critical data, add context and enhance the decision-making accuracy of AI models and systems.

<sup>3</sup> [https://www.informatica.com/lp/cdo-insights-2024-charting-a-course-to-ai-readiness\\_4656.html](https://www.informatica.com/lp/cdo-insights-2024-charting-a-course-to-ai-readiness_4656.html)

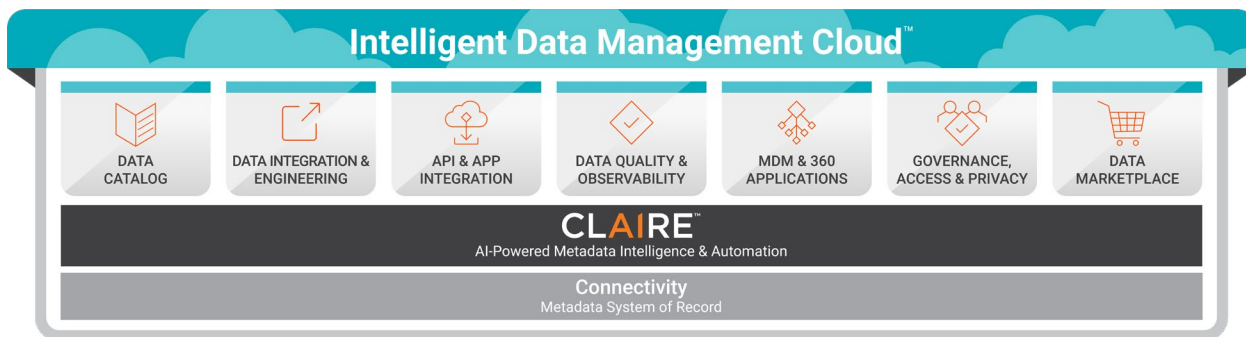


Figure 1. Informatica Intelligent Data Management Cloud provides you with a single, unified, AI-powered solution.

In addition, the Informatica **CLAIRE® copilot** helps improve performance by simplifying and automating data management processes, increasing efficiency, delivering trusted data faster and allowing teams to use AI responsibly. With **CLAIRE GPT**, a GenAI-powered data management assistant, companies can leverage a natural language-based chat interface that helps democratize business data access using natural language processing (NLP). It can also automate complex data management workflows to improve data literacy and drive greater productivity for data teams.

## Key Benefits

### Build Trust Among Stakeholders to Accelerate AI Adoption

Improve transparency into data operations and empower teams with automation underpinned by AI to derive accurate and reliable insights from AI. Enhancing the visibility and traceability of how data is collected, transformed and used within AI systems helps build trust among users, customers and regulators that AI models are trained on trustworthy and accurate data. Understanding how the AI model was developed, the training data used for creating it, its quality and lineage and the relevant policies applied helps to assure more predictable results when adopting AI applications.

### Improve the Reliability of AI Outcomes and Help Prevent Hallucinations

Monitor and remediate data issues to increase trust assurance and help ensure the availability of reliable data to fuel reliable AI outcomes. Comprehensive data profiling helps to continually assess the quality and integrity of data feeding into AI systems. It identifies data anomalies and potential biases that can mislead AI systems to fabricate or misrepresent information. Maintaining the integrity of **data pipelines** right through integration and transformation and their use in AI models reduces the likelihood of AI hallucinations by providing valid inputs.

## Reduce Compliance Risks and Create a Safer Environment for Innovation with AI

Apply Informatica's comprehensive, **AI-powered data governance** capabilities to help enforce consistent compliance frameworks and policy standards across large volumes of diverse and disparate data required for AI. IDMC's integrated capabilities such as data cataloging and governance, data masking, data privacy and protection, data marketplace and data access management provide companies with the agility to **navigate AI regulations** and help avoid costly fines. These capabilities support regulatory compliance and help reduce the privacy and security risks associated with data breaches and misuse. IDMC empowers users to apply AI confidently across their business initiatives to drive value.

## Help Ensure Data and AI Operations Are Accountable and Ethical

Informatica's metadata-driven predictive data intelligence solution helps ensure the data fueling AI systems is well-documented, managed and transparent throughout the data lifecycle. The ability to understand details about the data, such as structure, context, quality, interdependencies and sensitivity leads to informed data use and more reliable and explainable AI outcomes. Establishing clear accountability and structured processes to support responsible data use helps promptly and systematically identify and address deviations or risks associated with AI operations. IDMC's scalability, flexibility and automation help companies to deploy robust **data governance frameworks** that uphold ethical standards and align AI operations with the organization's values.

## Promote Secure Collaboration To Help Prevent Data and AI Misuse

Facilitate secure data sharing and empower data consumers with trusted data. Informatica Cloud Data Marketplace and **Cloud Data Access Management** capabilities allow organizations to create a centralized platform with self-service access to governed and protected data products that help meet regulatory requirements and risk tolerances. Effectively enforcing data access policies over sensitive data and monitoring movement, access and activity in line with responsible AI standards helps identify anomalous behaviors. This helps take quick and decisive actions that can minimize potential misuse.

## Key Capabilities

### Comprehensive Data Integration

Training complex AI models requires extensive volumes of data to minimize the risk of biases and improve decision-making accuracy. Informatica **Cloud Data Integration** service allows seamless out-of-the-box connectivity and flexibility to connect data from a broad range of data sources across cloud, on-premises and hybrid environments. It helps large-scale datasets spread throughout data silos to be integrated and made available for AI models at high speed through IDMC's scalable cloud infrastructure. Cloud Data Integration, a service of IDMC, helps ensure that the data used in AI models is cleansed, validated and transformed according to policies and guidelines for responsible use of data and AI. Cloud Data Integration allows continuous data pipeline integration and deployment (CI/CD) that helps AI models operate on relevant, up-to-date data, enhancing responsiveness and accuracy in dynamic environments. (Contact **Informatica** for the most current list of supported data sources.)

### Rapid Data Discovery and Cataloging

For AI systems to function fairly and accurately, it is critical to ensure the discoverability of data sets along with their context and quality. Informatica **Cloud Data Governance and Catalog** (CDGC) helps companies quickly find, understand and trust data assets to help ensure AI models are trained with a comprehensive and relevant data foundation. Enabling metadata-driven predictive data intelligence using a centralized "**catalog of catalogs**" for the enterprise allows data teams to gain comprehensive visibility of data across the enterprise landscape and utilize the most appropriate datasets to train AI models. By providing end-to-end **data lineage**, CDGC helps track where data originates, its attributes, its relationships and how it flows across systems. Understanding the source and journey of data provides transparency and accountability for data and AI models and helps ensure AI decisions are based on credible and trustworthy data.

### Automated Data and AI Governance and Policy Enforcement

Cloud Data Governance and Catalog (CDGC) empowers data teams to manage the data lifecycle in AI ecosystems, helping ensure that data remains secure and compliant with AI regulations and organizational policies. **CLAIRE** can automatically help organizations link policies to data assets by using AI-powered data classification upon discovery. This helps safeguard data from privacy and security risk exposure, helps ensure data is handled responsibly and improves collaboration by clearly defining data ownership and access permissions. Automated workflows within **CDGC** reduce the manual effort of curating, modifying and governing data assets in compliance with responsible AI principles and build trust among users and regulators. CDGC helps advance AI explainability by providing organizational visibility and transparency into models and their underlying algorithms. It also helps track and monitor model performance and key metrics, such as data drift, that may lead to unreliable business outcomes.



## Integrated Data Quality and Observability

Enabling AI systems to operate on accurate and unbiased data requires a comprehensive approach to data quality and observability. High-quality data is crucial for training reliable AI models to make fair and effective decisions. Informatica **Cloud Data Quality** empowers organizations to quickly identify, fix and monitor data quality issues across complex data environments. The IDMC service automatically performs data profiling to identify anomalies. It applies custom or out-of-the-box rules and data cleansing processes to support end-to-end data quality management and improvement. **Data observability** extends this by enabling ongoing monitoring and management of data health across the entire AI lifecycle. This continuous insight helps organizations detect and address data anomalies, drifts and biases proactively. By integrating robust data quality processes and observability tools, organizations can maintain transparency, accountability and trust in AI outcomes, ultimately fostering responsible AI practices crucial for ethical implementations.

## Master Data Management and 360 Applications

Standardized, cleansed, de-duplicated and enriched data across business-critical domains enables more accurate understanding of data and its interrelationships required for training and deploying reliable AI models and systems. Informatica **MDM and 360 Applications** offer a modern, all-in-one solution to help organizations create the trusted, contextual 360-degree data views needed for responsible AI. AI-powered automation and modern user interfaces increase productivity and facilitate better collaboration by aligning data across departments and making it readily accessible for analytics, AI and operations in a usable format. This helps ensure AI systems work with accurate, complete and up-to-date information, reducing the risks of errors or inconsistencies in AI-driven decisions. A consistent, authoritative source of truth for multiple data domains — including product, customer, supplier, employee, reference, financial and more — reduces discrepancies and potential biases in AI that are a result of incomplete and fragmented data management practices.

## Governed Data Sharing and Democratization

Ensuring accessibility to governed data assets for AI development across the organization is crucial for building inclusive AI models. Informatica **Cloud Data Marketplace** allows data owners to organize data into categories. It creates a place where data consumers can browse and shop for curated and validated data products compliant with AI policies and regulations. With detailed tracking of data usage and lineage available for each **data product** in the marketplace, data consumers can make informed decisions about requesting and sourcing those assets, thereby promoting transparency and accountability for handling data responsibly. As a central location for sharing and accessing data, **Cloud Data Marketplace** enhances collaboration between data producers and consumers, allowing companies to improve AI accuracy and reliability continuously through feedback and an exchange of knowledge.

## Secure Data Access Governance

Aligning data use with ethical guidelines, privacy standards and regulatory compliance reduces risks of data misuse, breaches and unauthorized access. Informatica **Cloud Data Access Management** (CDAM) helps companies ensure their data is safe and trustworthy when leveraged in AI systems while demonstrating adherence to relevant standards and regulations. CDAM features robust, policy-based access controls and data protection algorithms, including redaction, tokenization and generalization, to limit security and privacy risk exposure and promote appropriate data use. This helps prevent scenarios where AI might inadvertently discriminate or make unjustified conclusions based on improperly accessed or used data.

## Where data & AI come to



### Worldwide Headquarters

2100 Seaport Blvd., Redwood City, CA 94063, USA Phone: 650.385.5000, Toll-free in the US: 1.800.653.3871

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.

IN17-5021-0824

© Copyright Informatica LLC 2024. 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.

## For More Information

To learn more about modern data management tools that can help you achieve responsible AI outcomes, visit our website at [www.informatica.com](https://www.informatica.com).