Get smart about data and AI or get left behind

The essential buyer's guide to unified governance

Table of contents

Why unified governance is critical in the age of AI

The pace of AI innovation is creating an ever-widening chasm between what your organization needs to do with data and what you can realistically achieve. While the potential use cases for data have long outstripped most organizations' ability to support them, AI is now multiplying both the opportunities and risks.

The challenge isn't just about having more data or more use cases. It's about having the right foundation to scale data and AI initiatives safely. Organizations that try to move forward without unified governance will unleash chaos and risk. Those that build that foundation now will accelerate every data use case and achieve faster, better AI outcomes.

Key challenges in 2025:

- Fragmented governance tethers control and visibility to specific systems

- Data exists in pockets across applications, clouds and on-premises

- Business users are disconnected from technical systems

- AI multiplies the risks of unreliable and non-compliant use

The problem: Fragmented governance

The fundamental challenge you're facing isn't new. But AI has raised the stakes. Fragmented governance—where control and visibility are tethered to specific systems or platforms—creates a widening chasm between ambition and execution. This fragmentation prevents organizations from scaling data initiatives safely, precisely when they need to accelerate.

The root cause lies in how data landscapes have evolved: Organizations have accumulated data across disparate applications, multiple clouds and legacy systems. This creates not just technical silos, but strategic blind spots that impact:

- Data visibility and context

- Access control and compliance

- Business user empowerment

- Innovation velocity

As AI initiatives intensify across the enterprise, these challenges compound exponentially. Each new use case amplifies both the opportunity cost of fragmentation and the risks of unreliable or non-compliant data use.

7 out of 10 are internally developing apps that will incorporate AI or ML.

Source: Gartner, The Future of AI

The breakthrough: Unified governance powers the AI-ready enterprise

Collibra transforms this fragmented landscape through a fundamentally different approach to governance. Instead of reinforcing silos, our unified governance framework creates a comprehensive fabric of visibility and control that spans your entire data ecosystem. This breakthrough enables intelligent oversight throughout the complete data lifecycle—from ingestion through transformation to consumption.

The power lies in how unified governance weaves together:

* Automated intelligence that connects the right data with the right users

* Active linkages between datasets, policies and AI use cases

* Rich business context through an enterprise metadata graph

* Seamless collaboration between technical and business users

This isn't just governance modernized—it's governance reimagined as the foundation for accelerating innovation while ensuring safety at scale.

Data Confidence™: The new competitive advantage

The transformative impact of unified governance manifests in an organization's newfound ability to move with both speed and certainty. This unified approach dissolves the traditional tension between innovation and control, replacing it with a framework where safety and speed become mutually reinforcing forces.

Organizations with Data Confidence gain:

* First-mover advantages in AI adoption

* Accelerated data initiative deployment

* Automated guardrails that enable safe innovation

* Business user autonomy

When technical and business users can collaborate seamlessly within a trusted framework, innovation flows naturally from a foundation of high-quality data. That's Data Confidence—and it's becoming the defining characteristic of organizations that will lead in the AI era.

Your path forward

Organizations and data leaders face a pivotal choice: continue with fragmented governance and watch the data-AI gap widen, or unite governance to accelerate innovation safely. The work of unifying governance has to start now.

A unified governance platform empowers organizations to:

- Observe, track and protect data across all sources

- Automate documentation and traceability for AI use cases

- Enrich data context with every use through an enterprise metadata graph

- Bring technical and business users into collaborative data stewardship

The result? The power to scale data initiatives and accelerate new AI use cases with confidence.

How to use this guide

Whether you're just learning about unified governance or you're looking to refine your data management strategies, this buyer's guide is for you. It's structured to be both linear and modular. If you're a beginner, we recommend going through each section in order, building a step-by-step understanding. For those familiar with certain topics, feel free to jump to the sections that address your specific needs or challenges.

Who should use this guide

- **Chief data officers (CDOs):** As the top executives responsible for data management strategy, CDOs can use this guide to align strategy with business objectives

- **Data governance leads:** If you're tasked with initiating and operationalizing a data governance program, this guide offers a helpful path forward

- **IT leaders:** In the evolving IT landscape, the "shift left" movement is making data considerations crucial to the earliest stages of the development cycle. Use this guide to understand the role of unified governance in achieving better business outcomes

- **Line of business (LOB) leaders:** For those spearheading initiatives within specific departments — such as finance, marketing or operations — this guide offers specific examples of how unified governance can positively impact your organization

- **Data Scientists:** Data scientists need trusted data, especially with the volume of AI models that organizations are creating. Use this guide to understand how to work with data owners and offices and learn what specific capabilities will help you deliver trusted AI

- **Data Stewards:** As the guardians of data quality, data stewards get a comprehensive review of key considerations when choosing a solution for trusted data

Whether you're at the helm of your organization's data strategy or play a role in its execution, we hope this guide helps you accelerate your journey to unified governance.

The metadata (r)evolution

Metadata is at the heart of unified governance. What started as simple descriptors – file names, sizes and types — metadata is now leveraged for managing data lineage, data quality, data governance and much more.

Metadata serves multiple purposes including:

- **Search and discovery:** Like an advanced library catalog, metadata makes it easier to find the right data

- **Data governance:** Metadata provides insights into data ownership, usage policies and helps maintain data quality

- **Data lineage:** Metadata offers a clear picture of data's journey, revealing how it's been transformed, combined or processed over time

Active metadata: More than descriptions

Active metadata is designed to make metadata more actionable and integrated into data processes. Active metadata goes beyond static descriptions. It's dynamic, reflecting real-time changes in data. As you add new data sources or as data relationships change, active metadata adapts to offer the most current view of your data landscape.

Why is metadata so important? Here's why:

- **Informed decisions:** Make better decisions based on a more accurate and complete understanding of data

- **Efficiency:** Minimize time lost searching for the right dataset or trying to decipher what a data point means

- **Trust:** Know the lineage and quality of your data so you can be confident that you're working with reliable, up-to-date data

- **Compliance:** Especially in regulated industries, metadata helps ensure data handling and usage comply with relevant policies and regulations

Data cloud migration and unified governance

More than 85% of organizations will embrace cloud-first as a principle by 2025, according to Gartner.* And many will embark on a data cloud migration. The move to the cloud is often a catalyst for considering a unified governance solution. However, while the cloud promises compelling benefits from unlimited scale to seamless connectivity, the reality is more nuanced. Organizations are discovering that successful cloud adoption isn't just about lifting and shifting data—it's about transforming how data is governed so your technical and business users can trust, consume and comply with reliable, high-quality data.

Common challenges for migrations include:

- Prioritizing technical moves over business outcomes

- Fragmented governance that creates blind spots and control gaps

- Lift-and-shift approaches that result in cloud-based data swamps

- Siloed ownership models that limit strategic visibility and value

Rather than treating governance as an afterthought, leading organizations are using it as the foundation for successful cloud migration. A unified governance framework provides enterprise-wide visibility, active connections between business context and technical assets, automated controls that scale with cloud adoption and clear lineage from on-prem through cloud transformation.

This comprehensive approach ensures cloud migration becomes an opportunity to strengthen governance rather than fragment it even more. Organizations that embrace unified governance as part of their cloud strategy position themselves to realize the full transformative potential of cloud adoption while maintaining visibility and control.

Learn more about our proven framework for cloud migration success in our 'Data cloud migration' ebook.

Data cloud migration readiness checklist

Make sure your organization is prepared for data cloud migration. Our checklist covers the essential questions you should be asking before embarking on a data cloud migration journey.

Get the data cloud migration checklist

---

* Gartner, 'Gartner Says Cloud Will Be the Centerpiece of New Digital Experiences'

Getting smart about your unified governance decision

You've decided a governance solution is right for your organization. Now, how do you make the right decision to best serve your organizational needs? What should you expect to find in a solution that meets your current needs and can adapt to future growth?

The process of making a decision starts with thinking through the pros and cons of a platform versus a point solution. Although point solutions may seem sufficient, growing organizations often need platforms that can scale with increased data loads and more complex operations.

So how do I decide what choice is right? (Hint: It's a platform.)

Assess your immediate and future needs: While point solutions might seem adequate now, think about your growth trajectory and future requirements.

Consider integration capabilities: A platform offers seamless integration among its components, saving time and reducing complexities.

Do a cost-benefit analysis: Though platforms might look to be costlier initially, in the long run, they might prove to be more cost-effective than juggling multiple point solutions.

The pros and cons of platform versus point solutions

	Pros	Cons
Platform solution	• Comprehensive solution • Integrated components for seamless operation • Initial setup requires planning but is manageable with the right strategy • While comprehensive, it can scale to suit smaller operations • Potentially higher upfront investment balanced by long-term value	• Often easier to implement than a full platform • Might be better for small-scale operations
Point solution	• Tailored for specific needs • Scalable and can grow with an organization	• Can lead to fragmented data strategies • Integration challenges with other tools • Might require multiple vendors, leading to complexity

Building a data strategy that's right for you

When you're getting prepared to implement your governance solution, make sure you follow these critical steps to ensure you make the right decisions.

- Understand your business goals: Align your data strategy with what your organization aims to achieve

- Assess data maturity: Know where you stand in terms of data infrastructure and what gaps you need to fill

- Stakeholder collaboration: Engage with all relevant stakeholders, from IT to business leaders, to get a comprehensive understanding of needs

An integrated fabric: The unified governance platform

Unified governance isn't just a collection of tools—it's an integrated fabric that enables organizations to move with confidence and speed. The key components work in concert to create a foundation that accelerates innovation while ensuring safety at scale.

- Data catalog: An organized inventory of data assets that empowers stakeholders to find and use the right data

- Data governance: The business glossary and data stewardship tools needed to support organizational practices that define data ownership and ensure trust in your data

- AI governance: The application of rules, processes and responsibilities to drive maximum value by ensuring streamlined, ethical AI practices that mitigate risk and protect privacy

- Data privacy: A tool to ensure sensitive data is protected and compliance with data protection regulations is maintained

- Policy control center: A centralized location to define, enforce, and monitor data access and usage policies across the organization

- Data lineage: A visual representation of data flow, helping to manage the origin, transformation and consumption of data

- Data quality & observability: A solution for achieving high-quality, accurate and consistent data as well as insights into data health

By understanding the components of a unified governance platform and aligning your organization's needs with what's available in the market, you can make an informed decision on the right solution.

Collibra

Data catalog: The foundation of unified governance

A data catalog creates and maintains an inventory of your organization's data assets across your entire digital environment. It enables the discovery, description and organization of data assets. Data catalogs do this by leveraging metadata.

Most importantly, the modern data catalog serves as the architectural foundation for unified governance, providing the crucial layer for joining technical capabilities and business value. By transforming raw metadata into actionable intelligence, data catalogs create a dynamic knowledge network that grows more valuable with every interaction.

Why is a data catalog important?

- Delivers end-to-end visibility: A data catalog provides a comprehensive view of your data assets, showing where they come from, how they transform and where they're consumed as well as providing useful business context

- Maximizes the value of your data: Your biggest opportunity is in the data you already manage. A data catalog helps organizations go from drowning in data to charting a path to unified governance Use cases

Choosing a data catalog
Look for these capabilities when considering a data catalog:

- **Connectors:** Make sure your data catalog offers a centralized repository of connectors to your most widely and actively used data sources. This ensures your catalog is as comprehensive as possible

- **Curation:** Make sure your data catalog offers tagging, grouping and documentation: to improve data discoverability and understanding. Tagging can help users quickly find related datasets, grouping can categorize data and documentation provides context and clarity

- **Automated classification:** Automation accelerates and drives efficiencies in classifying ingested data, eliminating error-prone manual efforts and enhancing accuracy

- **Collaboration tools:** An important element for operationalizing your data catalog is the capacity for users to communicate and interact with each other within the data context by sharing comments, annotations and shared insights. This can foster a more collaborative data culture, which is key for adoption

- **Data marketplace:** A distinguishing capability for enterprise data catalogs, a data marketplace is a platform where users at all levels can 'shop' for datasets easily, helping them discover, share and collaborate around data

Use case by role

Marketing analytics: A marketer can use the catalog to find reports related to customer behavior, leading to more informed campaign strategies.

Supply chain optimization: Managers can locate data on suppliers, inventory levels, and transportation and build dashboards, leading to a more efficient supply chain.

Data science: Data scientists looking for datasets to use as training data for an AI model can leverage a data catalog to find relevant data.

Use case by industry

Healthcare: Hospitals and healthcare providers can use catalogs to track patient data, treatment histories and outcomes for improved patient care.

Finance: Banks and financial institutions can track transactions, customer interactions and market data to build AI-based predictive analytics leading to better risk management and customer service,

Government: Agencies can streamline the management of vast public records, creating comprehensive repositories that organize the inventory of datasets, track usage and secure access.

Data lineage: Map the data journey

In an era where data powers every strategic decision and AI initiative, data lineage has evolved from a technical necessity to a critical business capability. A data lineage solution maps and illuminates the complex pathways through which data flows, transforms and creates value across your enterprise.

Within a unified governance framework, data lineage creates the intelligence fabric that connects your entire data ecosystem. Every data transformation becomes a signal that enriches your understanding of data health and utilization patterns. Real-time visibility into data relationships enables confident decision-making at scale, while automated impact analysis accelerates the pace of innovation.

As data flows through your organization, lineage transforms raw movement into actionable insights that drive strategic value.

Why is data lineage important?

Data doesn't exist in isolation; it flows, transforms, gets consumed and influences decisions. Knowing its journey is akin to understanding its story.

- **Trust and accuracy:** Just as you'd trust a story more if you knew its source, tracing data's lineage boosts confidence in its accuracy

- **Data quality:** If data seems incorrect or inconsistent, lineage helps pinpoint the stage where the discrepancy occurred

- **Compliance and audits:** Regulations often require businesses to explain their data sources and transformations. Data lineage offers a clear map for compliance teams and auditors

Key features to look for

- **Automated lineage discovery:** Manual tracking can be tedious and prone to errors. Automation ensures a thorough and efficient tracing of data flow

- **Visualization capabilities:** A graphical representation of data flow is more intuitive and actionable than textual logs. Visualizations help stakeholders understand and navigate the data's journey

- **Integration with a wide variety of data sources:** Data can reside in databases, cloud storage, third-party applications, and more. A good lineage solution should seamlessly integrate with diverse data sources

- **Support for technical and business lineage:** While technical lineage focuses on the nitty-gritty of data flow, transformations, and systems, business lineage provides a higher-level view, mapping data's impact on business processes and decisions

- **Ability to handle large volumes of data:** Modern enterprises deal with vast amounts of data. The solution should scale without compromising performance

Use cases by use case:

- **Regulatory compliance:** Demonstrate compliance — for example, to GDPR — by mapping the journey of personal data through various systems and processes

- **Data migration:** Assess the potential impact of planned data changes to business units across the organization, minimizing disruptions

- **Data lineage:** Track and document the lineage of input, training and output data of AI models, ensuring transparency and accountability

Use cases by industry:

- **Insurance:** Tracking underwriting data history, claim records, and customer interactions, ensuring regulatory compliance and assisting in decision-making

- **Telecom:** Troubleshooting network issues, optimizing resources and managing vast amounts of data from call records, network traffic and customer information

- **Government:** Operationalizing a transparent trail of accountability, crucial for policy analysis and maintaining public trust Data governance: Ensure collaboration and compliance

Data governance: Ensure collaboration and compliance

Data governance is the practice of managing and organizing data and processes to enable collaboration and compliant access to data. Data governance allows your data citizens — and that's everyone in your organization — to create value from data assets. Leveraging a business glossary and a data dictionary as well as stewardship management and reference data management, enterprise data governance provides a single location to find, understand and create a shared language around data.

Unfortunately, when data governance goes badly, it can lead to costly consequences, including:

- **Misinformation:** Incorrect decisions based on inaccurate data

- **Security breaches:** Vulnerabilities due to improper data handling

- **Regulatory violations:** Potential legal consequences and hefty fines

- **Wasted resources:** Time and money spent rectifying data errors

Increasing regulations across the globe

With data breaches making headlines, nations are intensifying their data protection regulations. Examples include GDPR in Europe, CCPA in California and even AI specific laws like the EU AI Act which has data specific requirements. Organizations must adapt to these changing landscapes to avoid penalties and maintain customer trust.

Key features to look for

- **Comprehensive business glossary:** A centralized location for all data-related terminologies ensuring consistency in understanding and usage

- **Stewardship management:** Identifying owners for the data who ensure data quality and usage adheres to organizational policies

- **Reference data management:** Managing standard data assets that are repeatedly used across the organization

- **Centralized policy management:** A singular point to define, manage and monitor data-related policies

- **Workflows:** Automated processes for data approval, quality checks and updates

- **Flexible operating model:** A flexible model allows you to design your environment to meet your specific needs

Use cases by use case:

- **Mergers and acquisitions:** Ensuring data consistency and quality post-merger

- **Customer analytics:** Providing accurate insights based on high-quality customer data

- **Product development:** Utilizing data to inform product innovation and address market needs effectively

Use cases by industry:

- **Manufacturing:** Oversee supply chain data, production quality and inventory management

- **Agriculture:** Streamline management of crop yield data, soil health records and climate impact assessments

- **Utilities:** Manage infrastructure data, consumption analytics and green energy initiatives

AI governance: AI starts with trusted data

It's clear now that the introduction of ChatGPT in late 2022 was a breakthrough moment in the history of technology and, perhaps, the evolution of society itself. Around the world, enterprises are now fiercely competing to develop AI applications.

To stay competitive, you need to ramp up your AI initiatives — and launch applications that leverage the groundbreaking capabilities of LLM technology. But successfully integrating AI initiatives into your roadmap will require a rigorous approach. To ensure AI is used responsibly, forward-thinking enterprises are leveraging AI governance.

AI governance is the application of rules, processes and responsibilities to drive maximum value from your automated data products by ensuring applicable, streamlined and ethical AI practices that mitigate risk and protect privacy.

Why AI governance is important

The truth is data is the backbone of AI, and if the data is bad, the AI models trained on it will produce results that look good but are fundamentally flawed. The implications for companies building AI applications are profound. And as AI's impact grows, so does the scrutiny. The European Union's AI Act is an example of regulations emerging to ensure ethical AI use. Non-compliance can lead to:

- Legal repercussions

- Loss of customer trust

- Brand damage

Company policies and government regulations

Both organizational policies and governmental regulations work hand in hand. While government regulations set the standard, company policies can further refine and detail AI usage specific to organizational needs.

Key features to look for

A well-defined AI governance solution is built on a framework that should encompass:

- Clear AI objectives and principles

- Data quality and integrity checks

- Model transparency and interpretability measures

- Regular audits and updates

- Ethical considerations, including fairness and bias checks

- Legal and data compliance

As data becomes increasingly critical to organizational decision-making, the emphasis on data governance becomes more and more business-critical. Whether you're looking at traditional data governance or the emerging practice of AI governance, the presence of a clear, comprehensive framework is essential in today's data-driven world.

Use cases by use case:

- **Model lifecycle management:** Accelerating AI development through automated quality controls and documentation flows

- **Enterprise risk management:** Ensuring model transparency and regulatory compliance across AI initiatives

- **Innovation pipeline:** Enabling rapid experimentation while maintaining rigorous governance standards

Use cases by industry:

- **Financial services:** Powering trusted AI adoption across risk modeling, fraud detection, and personalization

- **Healthcare:** Enabling compliant AI innovation in diagnostics, care optimization, and clinical workflows

- **Manufacturing:** Accelerating smart factory initiatives while maintaining complete model visibility

Data privacy: Handle data properly

As organizations race to harness the transformative power of AI and advanced analytics, data privacy has evolved from a compliance checkbox to a strategic differentiator.

Where data powers competitive advantage, the ability to protect and ethically leverage sensitive information becomes fundamental to maintaining stakeholder trust and enabling sustainable innovation.

The evolution of personal data

Today's digital landscape has fundamentally transformed what constitutes personal information. Beyond traditional identifiers, the modern privacy perimeter must account for:

- Digital fingerprints that span device signatures to behavioral patterns

- Biometric markers that uniquely identify individuals

- AI-generated insights that can reconstruct identity from seemingly anonymous data

- Cross-platform digital footprints that create unprecedented visibility into personal lives

In the news: The 2023 X data breach

In January 2023, it was widely reported that an X (formerly Twitter) data leak exposed email addresses linked to more than 200 million profiles. Stemming from an API flaw, the breach put X's pseudonymous users at risk by potentially unmasking their real identities as well as fueling phishing and identity theft.

Key features to look for

- **Automated sensitive data discovery:** Tools that automatically scan and identify sensitive data, ensuring that PII and other private data are consistently flagged and protected

- **Dynamic data mapping:** Tools that offer real-time mapping of data sources and their relationships, ensuring that data flows can be quickly analyzed for potential privacy risks

- **Embedded privacy by design:** Systems that are built from the ground up with privacy in mind, ensuring that data protection is not an afterthought but a core feature

- **Persona-based UI:** User interfaces tailored to different user roles, ensuring that users only access and view data pertinent to their tasks, reducing the risk of unintentional data exposure

Use cases by use case

- **Customer service:** Ensuring customer service reps see only the necessary data to assist a customer, not their entire profile

- **Product development:** Leveraging customer feedback and usage data to inform the product roadmap

- **Legal:** Managing global data protection laws by automating the identification and treatment of sensitive data

Use cases by industry

Technology: Ensure PII isn't utilized in AI models (if it's not required), protecting user privacy while harnessing data for better service and targeted marketing.

- E-commerce: Keep customer data, browsing history and personal preferences safe

- Education: Safeguard student records, including academic performance and personal information, while facilitating educational needs

Data privacy is not just a legal necessity but also a marker of trust and reputation for organizations. As individuals, understanding and controlling our data footprint is the first step in ensuring our privacy in an increasingly interconnected world.

Define and control data access

Data access governance is the process of managing and controlling access to organizational data.

Establishing comprehensive data access policies is a cornerstone of effective data governance and security, particularly in industries where sensitive information is routinely handled. These policies define who has the right to view or manipulate data within an organization, ensuring employees only have access to the information necessary for their roles.

The goals of a scalable data access governance program are to enhance data security and regulatory compliance, as well as prevent unauthorized access and data breaches.

Why is policy control important?

No organization wants the notoriety that follows a data breach or an example of an AI model that uses data it shouldn't. Policy control is paramount, serving as the framework for dictating how data can be accessed, shared and used. When you can control data access, your organization can:

- Mitigate risk

- Ensure compliance

- Build trust

Key features to look for

- **Data access policy overview:** Govern data policies from a single pane of glass

- **No-code policy builder:** Easily write and push policies to the cloud

- **Automatic data classification:** Take advantage of automatic classification to understand the content and sensitivity of data.

- **Enterprise data catalog:** Organize and maintain an inventory of your organization's sensitive data across departments and domains.

- **Granular access controls and protection:** Enable access controls at the column and row level

Use cases by use case

- **R&D:** Ensure the right teams can access proprietary and customer data while protecting confidential information from access

- **HR analytics:** Analyze workforce data for talent management and strategic hiring plans while safeguarding personal data

- **Auditing and reporting:** Facilitate auditors' access to financial records and operational data without exposing sensitive business or employee information

Use cases byy industry

- **Automotive:** Ensuring auto manufacturers both safeguard personal driver information while using data to innovate and improve the driving experience and supporting infrastructure

- **Entertainment:** Leveraging user preferences for personalized experiences while maintaining confidentiality of usage habits and subscription details

- **Local government:** Protecting citizen data from unauthorized access while ensuring it is available for public services, policy making and intra-agency collaboration

Data quality & observability: Ensure value and transparency

Modern stacks are complicated. While amassing vast amounts of data is an achievement, it doesn't automatically translate to value. And data without quality is like a recipe without any good ingredients — there's potentially a meal at the end of the process, but if it's not going to taste good, the recipe isn't very useful.

Data quality refers to the ability of data to fit its intended purpose, ensuring it is accurate, complete and reliable for decision-making. Observability in the context of data is the ability to fully understand the health, status, and performance of data systems and pipelines through monitoring, logging and tracing. Together, data quality and observability empower organizations to trust their data-driven decisions by providing transparency and insights into your data operations.

$15M

The average annual financial repercussions of poor data quality.

Source: Gartner Data Quality Market Survey

Key features to look for

- **Data profiling:** Understand your data's structure, content and quality. This is the preliminary step in identifying areas that need attention

- **Data cleansing:** Clean or remove corrupt, inaccurate or erroneous bits of data to maintain its integrity and reliability

- **Data enrichment:** Enhance the data's value by appending related information from external sources. This not only adds depth but also context

- **Data monitoring:** Keep an eagle eye on data streams to detect anomalies, inconsistencies or any deviation from the defined quality standards

- **Data pushdown:** Sick of being blindsided with egress fees? Leveraging tools that address data quality at the source can amplify security, hasten processing and maximize ROI

Use cases by use case

- E-commerce: Ensure personalized recommendations that resonate with users, leading to better conversion rates and a better user experience

- Risk management: Identify potential risks and take preemptive mitigation measures

- Operational efficiency: Drive accuracy of operational metrics, improved process efficiencies and reduced downtime

Use cases by industry

- **Healthcare:** Providing correct diagnoses, treatment plans and care outcomes. Ensuring data quality can literally be a matter of life and death in this sector

- **Energy:** Managing resources, forecasting demand and managing the grid to ensure reliability and efficiency

- **Retail:** Driving more effective inventory planning and more personalized shopping experiences

Unveiling the power of active metadata

Active metadata is a dynamic, interconnected representation of data. Active metadata can span a wide range of asset types, weaving together information about each asset. The integrated view provides rich content and context to data, ensuring it can be both trusted and acted upon. Unlike static metadata which offers a snapshot, active metadata is constantly evolving, reflecting real-time changes and integrations.

Key benefits of active metadata

- Generate greater visibility into the data landscape: With automatic data classification and auto-linking of data sets, business terms, policies, processes and more, data curators and data consumers can collaborate on business semantics for trusted data

- Evaluate the right data for your needs: Data profiling, data scoring and crowdsourced ratings and reviews strengthen data context and allow business analysts and data scientists to evaluate and choose the best data for their purposes

- Enhance data shopping experience: With highly relevant and rich business context around data, users have a more intuitive and simplified data shopping experience, which provides the right data with the right context to the right users

- Deliver faster insights: Automated discovery, understanding and collaborative data access for business analysts and data scientists reduce time to insights

Active metadata isn't just another buzzword. It represents a shift toward a more interconnected, dynamic and actionable data environment. As data continues to grow both in volume and complexity, tools and methodologies like these will be instrumental in ensuring that data remains an asset, not a challenge.

Collibra Platform success stories

There's no shortage of organizations across virtually every industry that's benefited from Collibra. These organizations use a unified governance platform to do more with trusted data.

U.S. Department *of* Defense

The Office of the Secretary of Defense

The problem:

The Office of the Secretary of Defense (OSD) and its Comptroller Office of the Undersecretary of Defense are responsible for US defense policy, planning, resource management and program evaluation. In carrying out those duties, the OSD's day-to-day decisions impact a wide range of operations, including human resources, weapons acquisition, research, intelligence and fiscal policy across all of the US armed forces (Army, Navy, Marine Corps and Air Force). Steering an organization of that size means data is crucial to its decision-making processes. However, ensuring decision makers have access to the right data, can trust in its accuracy and understand its context is by no means a simple task.

The solution:

Recognizing the challenge, Greg Little, Director, CFO Data Transformation Office, partnered with Collibra to launch their Advana (Advancing Analytics) Data Catalog to create a centralized platform to support data and analytics across the organization. The Advana program is still relatively nascent, but has already made significant strides in its mission, with the team tracking key performance indicators to ensure they remain on the right path and can measure their progress.

Learn how the Office of the Secretary of Defense unified their governance strategy and accelerated their data and analytics use cases.

FINANCIAL SERVICES

Equifax

The problem:

A global data analytics and technology company, Equifax plays a critical role helping consumers and companies make critical decisions with greater confidence. Their unique blend of differentiated data, analytics, and technology help them create the insights that power organizations to move forward in their complex, competitive markets. Equifax is central to the process of applying for a job, buying a car, getting a home loan, and many other common financial transactions. However, organizational silos were limiting the company's ability to empower stakeholders. With data at the center of Equifax's business — and more than 50,000 data sources — digital transformation was their most important initiative.

The solution:

Equifax set out to create a seamless, centralized data fabric platform that would optimize the data pipeline and empower stakeholders to find and access data. To combat siloed data restricted by local rules and regulations, Collibra offered a central data platform accessible by stakeholders across five continents — from North America to Latin America, Asia to Australia. The intuitive, easy-to-use Collibra experience has enabled Equifax professionals from data scientists to marketing managers to get more value out of data.

Learn how Equifax de-fragged their data governance strategy so that their entire organization could access the data.
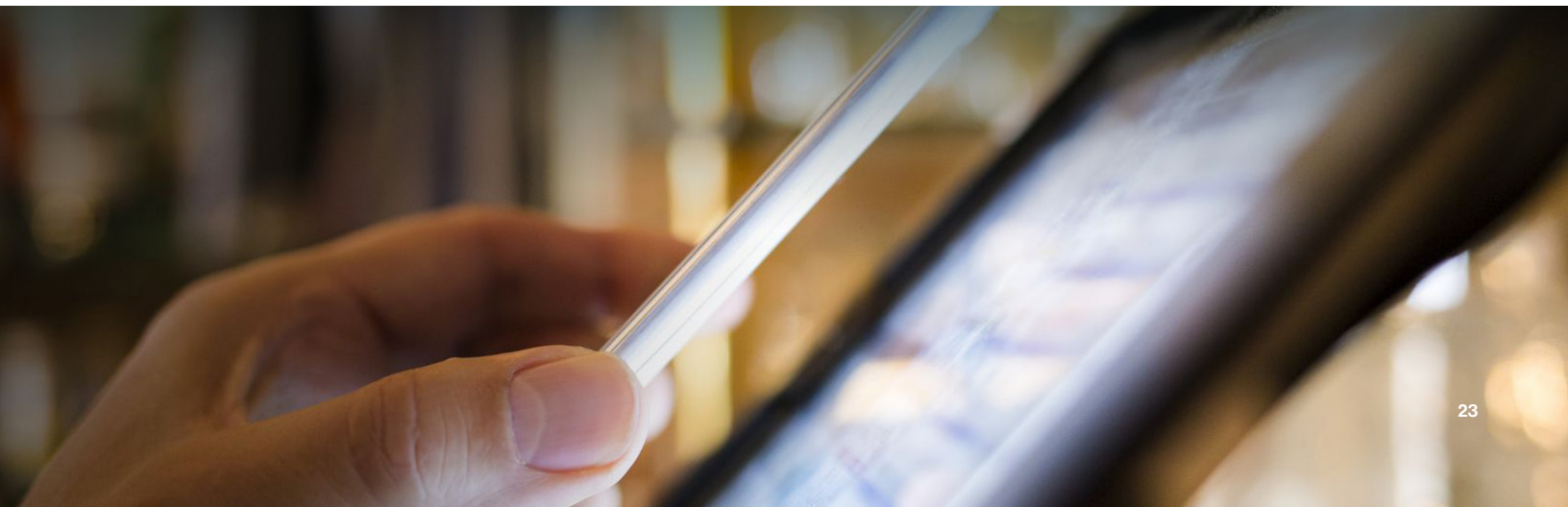
HEALTHCARE

Envision Healthcare

The problem:

Insurance claim denials are a growing problem in U.S. healthcare. For the national hospital-based physician's group Envision Healthcare — an organization that provides services to 30 million U.S. patients every year — data anomalies can lead to denials. But with more than 50 data collection systems, the company struggled to drive data governance and data quality. And the consequences were being felt all the way to the bottom line.

The solution:

Envision implemented Collibra — and utilized Collibra Rapid Start Implementation Service to drive deployment. Today, Envision uses Collibra Platform running on AWS, including Data Quality & Observability, Data Catalog, Data Governance and Data Lineage capabilities. Collibra is integrated with several applications, including Microsoft SQL Server and Tableau. Envision uses Data Catalog and Data Quality & Observability to ingest data warehouse documentation, understand the data and develop an up-to-date business glossary. This enables Envision data citizens to search, find and interpret data. Adoption is robust. Approximately 900 people are accessing Collibra, and 60 data stewards have been appointed.

Learn how Envision Healthcare ensures Data Confidence across their organization.

RETAIL/CONSUMER GOODS

HEINEKEN

The problem:

One of the most recognizable beer brands in the world, HEINEKEN recently launched its EverGreen business strategy to drive growth and sustainability, become the best-connected brewer, and uphold the HEINEKEN quality brand experience. But the company needed better control and understanding of the vast amount of data it generates from multiple sources — such as customer information, taste and buying trends, brewing process sensors and the supply chain.

The solution:

HEINEKEN wanted a solution that would enable them to develop and implement effective data governance across the organization and give employees the visibility and access to data needed to make more impactful business decisions. Today, Collibra acts as a centralized source of truth for the business. All of HEINEKEN's 85,000 employees have access to Collibra, with each OpCo having their important data and reports at their fingertips. "Before Collibra, HEINEKEN was flying in the dark," commented de Geus, "Now we can shine a light on our data and make our data landscape visible. If an employee wants to know something, they go to Collibra. We are making Collibra a one-stop-shop and single source of truth for all things data."

Learn how HEINEKEN is accelerating their data and AI use cases.

LIFE SCIENCES

Daiichi Sankyo Europe

The problem:

Daiichi Sankyo Europe — a subsidiary of Daiichi Sankyo Group — discovers, develops and delivers innovative medicines for people with cancer, cardiovascular and other diseases with high unmet medical needs. Operating in 20 countries, Daiichi Sankyo Europe has more than 17,000 staff worldwide and over 120 years of scientific expertise. Daiichi Sankyo Europe sought a way to use data to gain a 360-degree understanding of the healthcare professional community, their needs, how best to interact with and serve them, and to predict future products and offers.

The solution:

Today, Daiichi Sankyo Europe employs a cloud-first strategy using a Collibra Platform that includes Data Catalog and Data Governance as the foundation of the company's data governance platform. Collibra integrates with Snowflake, Power BI, a CRM application for the life sciences industry, and dbt, a data transformation tool. Because Daiichi Sankyo Europe has 15 country affiliates, they adopted a hybrid data mesh model that harmonizes the dynamics of decentralized data governance. "We are moving towards a world where data isn't a privilege but a fundamental right — an ecosystem where every team and individual has access to the data they need to drive innovation and make informed decisions," explained Pavel.

Learn how Daiichi Sankyo Europe centralized their data efforts to create a unified data governance strategy.

EDUCATION

# Western Governors University

The problem:

Western Governors University (or WGU) was established in 1997 to expand access to high-quality, affordable higher education and serves 140,000+ students nationwide. Formerly, the university had struggled to make the best use of its data. "The problem was we could not get our arms around the data,". "We could not find what we needed or what information we had. When we did find it, it was difficult to understand and we did not know if it was reliable," said Garth Gehlbach, Head of Data Governance at WGU. Data was largely confined to departmental silos and there were limited enterprise-wide views. Finding data to create a report and perform business analysis took a long time and often took several iterations to get it right. WGU had some great data assets. They just needed unlocking and to help achieve that, he needed the right solution.

The solution:

WGU is using Collibra Platform as the foundation of its Data Governance Center. University departments have submitted over 1,400 business glossary terms and Collibra now holds over 850,000 assets, which represent technical schemas, business and governance context, and relationships. Today, Collibra is fulfilling its purpose of being the system of truth for enterprise metadata and providing a robust Data Context Graph that enables WGU data citizens to find, understand, trust, and curate data.

Learn why Collibra is the system of truth for Western Governors University.
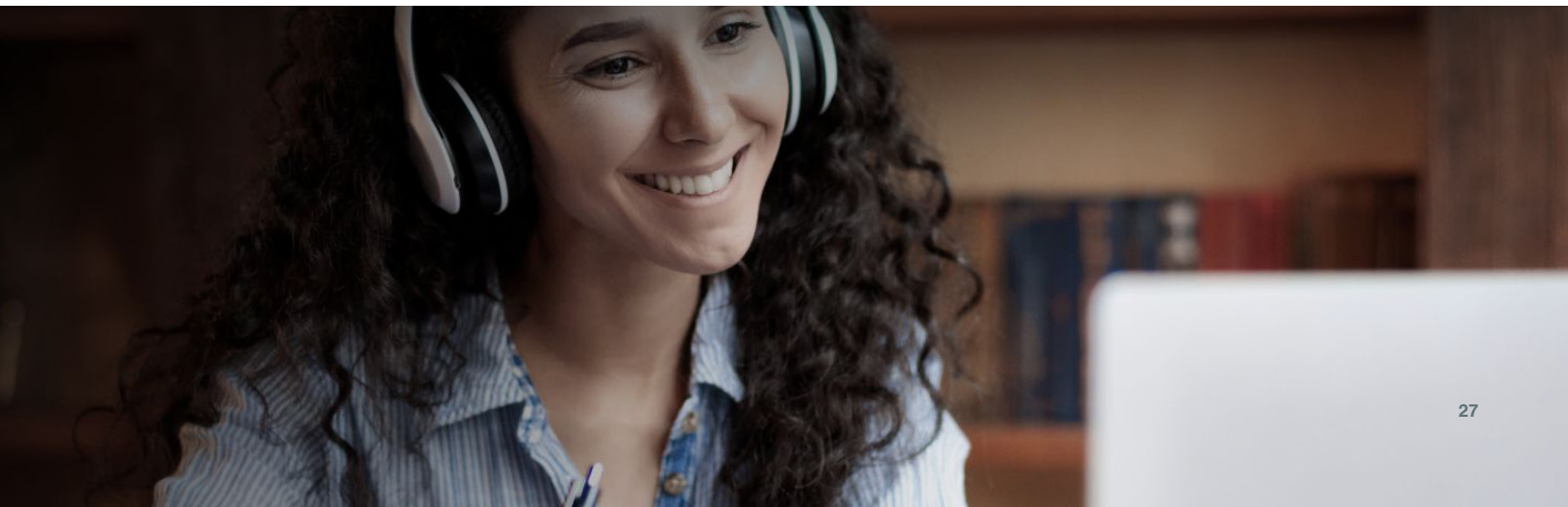
Adoption: To start small or move fast?

Implementing a unified governance solution is a strategic initiative that allows an organization to harness the full value of its data. When it comes to adoption, you'll want to create an implementation plan that's right for your organization.

Starting small

There are a few different ways to get started; the right one for your organization should align with your objectives. When you're ready, you can launch your initiative by starting with one (or more) of these critical areas:

- Data catalog

- Data quality

- Data governance

A common starting point is to establish a data catalog. As described earlier, a data catalog is a centralized repository where your metadata is stored. This simplicity and accessibility can drive better compliance with data governance and quality standards from the outset.

Implementing data quality measures early is another critical step and potential starting point. It can be as straightforward as defining data quality rules, metrics and setting up processes for regular audits. Ensuring early on that data is clean, consistent and trustworthy can help avoid the costly, complex process of fixing issues later.

Governance frameworks are also foundational — establishing the policies, procedures, and standards for data usage and handling across the organization.

Moving fast

But maybe you need to move fast? To do so, consider starting with a single use case and then using agile methodologies. This involves short sprints that focus on delivering small chunks of functionality quickly. It allows the business to see value early and provides an opportunity for feedback that can guide subsequent iterations and future use cases.

Automation tools can play a significant role in a rapid start. For example, automated data discovery and classification can accelerate the creation of a data catalog. Similarly, automated quality checks can be implemented to ensure new data sources meet quality thresholds before they are ingested into the system.

Expansion and scalability

Once you have your foundations in place, you can gradually expand the scope of your data solution. The key is to maintain flexibility and adaptability in the system to handle diverse data types and sources. As your organization matures in its use of data, observability becomes crucial. It will enable the team to monitor the system's health and performance, understand data lineage, and quickly diagnose and resolve issues.

Continuous improvement

Finally, a successful implementation is not a one-time effort but requires continuous improvement. Regular reviews of data quality, governance policies and the overall data strategy helps ensure they align with the organization's objectives.

By keeping the initial steps manageable and rapidly delivering value, you can build momentum for your data initiatives, securing buy-in from stakeholders and laying the groundwork for a data-driven culture.

Often, an experienced, trusted integration partner can help you implement a data solution that reflects industry best practices, a roadmap to provide direction and a rollout strategy that can drive adoption.

The unified governance checklist

We've covered a lot of ground in these pages. For good reason. There's a lot that's going into your decision about a unified governance platform. So we've summarized everything into a streamlined checklist. Keep it with you to ensure you're asking the right questions as you evaluate platforms.

- Does it offer a data catalog that organizes your data assets in a way that makes them easy to find and use?

- Does it provide data governance that includes a business glossary and data stewardship tools that effectively define data ownership and build trust in your data?

- Does it help accelerate your AI initiatives with AI governance that reduces risk, protects privacy and drives trusted AI?

- Does it include robust data privacy tools to protect sensitive data and comply with data protection regulations?

- Does it have a central policy control system to define, enforce and monitor data access and usage policies across your organization?

- Does it offer data lineage to help effectively manage and visualize the origin, transformation and consumption of data?

- Does it come with a data quality & observability solution to drive data quality, accuracy and consistency?

Unifying governance for organizational success

Accelerating all your data and AI use cases

The gap between what your organization needs to do with data and what you can accomplish continues to widen, especially as AI adoption accelerates. Traditional fragmented governance, with controls tethered to specific systems and sources, prevents agencies from scaling their data and AI initiatives safely.

The Collibra Platform—which provides several deployment models including FedRAMP authorized cloud and self-hosted—offers a different approach through unified governance. By untethering governance from systems and sources, we give agencies true visibility, context and control throughout the full data cycle. Our platform enables automated access control and comprehensive AI governance, creating active links between datasets, policies and AI use cases.

The result? Data Confidence. Now you can accelerate all your data and AI use cases, safely.

Are you ready to transform how your organization manages data and AI? Learn more about Collibra.