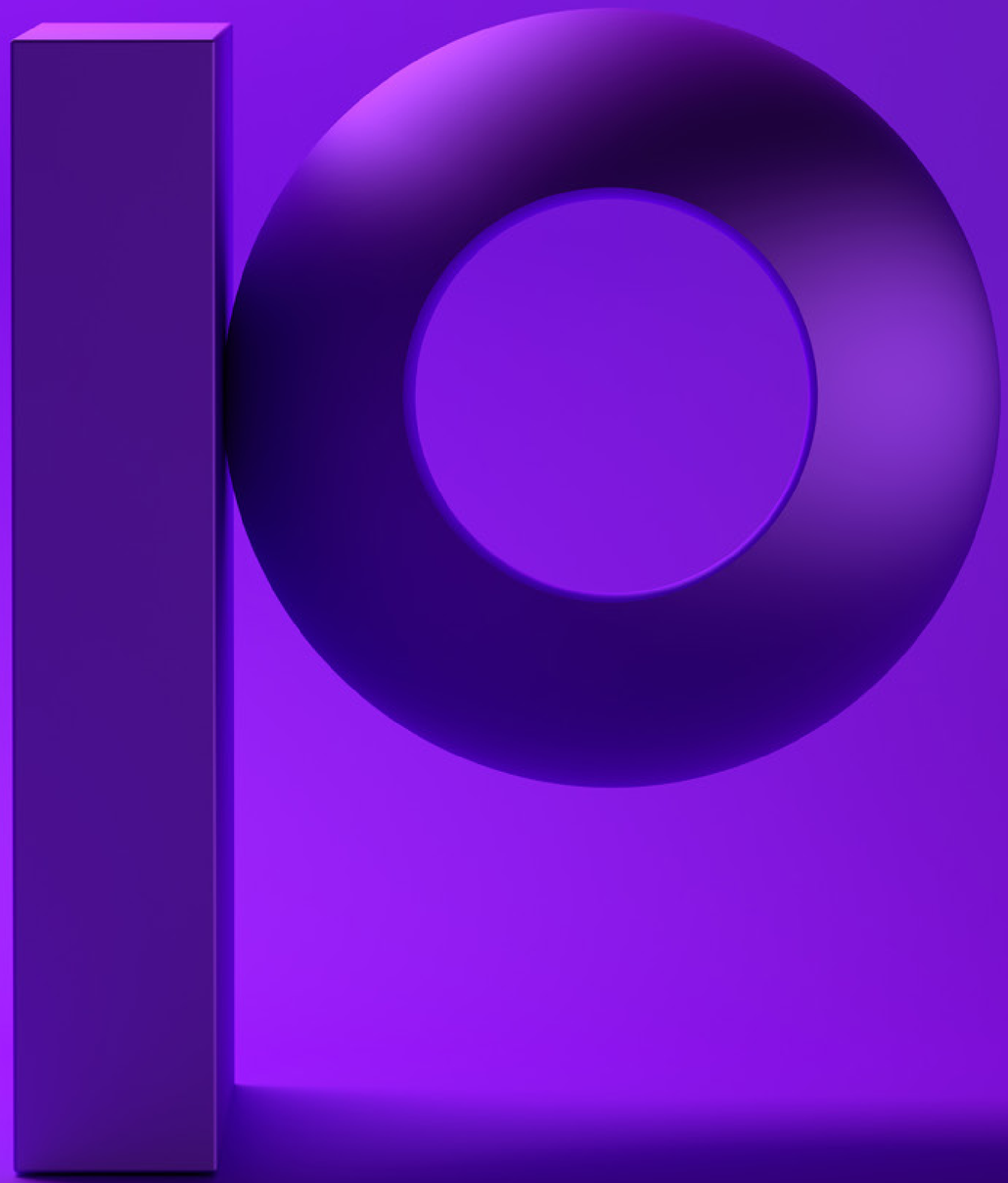
precisely

Data Governance 101:

Back to Basics to Support Evolving Analytics & AI Ecosystems



Mission impossible? Overcoming data governance challenges

Most organizations today already have a data governance framework in place — but many are realizing that what worked a few years ago isn't enough for today's data-driven and AI-powered world. As enterprises depend more on data and analytics to drive business intelligence, innovation, and automation, governance programs must evolve to deliver measurable value. AI initiatives, in particular, raise the stakes: models trained on incomplete, biased, or poorly defined data can generate unreliable results and expose the organization to ethical or regulatory risk.

Strong data governance ensures that the data fueling both analytics and AI systems is accurate, consistent, and transparent, giving leadership confidence that insights can be trusted. Even mature programs benefit from revisiting the fundamentals. Can your organization clearly answer: "Am I using the right data?" and "Can I trust the quality of my data?" These basic questions reveal whether your governance framework is optimized for today's challenges. The following six questions offer a practical way to assess and strengthen your approach, ensuring your governance program continues to deliver value as data ecosystems and business needs evolve.

6 Key questions to help identify the strength of your data governance program

1. Why is this data important?
2. Does it mean the same thing to everyone?
3. Where does the data come from?
4. Can I trust it?
5. How do I find it?
6. Who do I ask?

Why is this data important?

No matter how established your governance framework is, it's worth re-examining whether it's focused on the data that truly matters. Not every piece of data in your organization needs the same level of attention. Effective data governance begins and continually improves by identifying the data that truly matters: the information that drives business initiatives, supports KPIs, and includes personally identifiable information (PII).

By refocusing governance on this high-impact data, you create a foundation for accurate reporting and trustworthy insights. Teams can clearly show how data contributes to business results and confidently demonstrate its value. Spreading governance efforts across every dataset only adds complexity without delivering meaningful outcomes.

Strong governance of critical data also improves security and compliance. Understanding where sensitive data lives, who uses it, and how it moves across systems allows you to define and enforce the right policies for protection and access. It also ensures that the metrics driving key decisions are accurate, consistent, and complete.

AI adds a new dimension to this challenge. As organizations increasingly rely on AI and machine learning models, the data feeding those models must be governed just as carefully as traditional data sources. Models — and the data behind them — constantly evolve. Without clear oversight, even small shifts in data quality or meaning can erode model performance, trust, and business value.

A refocused governance strategy helps you manage what matters most. By prioritizing the data that drives business outcomes and fuels AI models, organizations reduce risk, improve accountability, and build long-term confidence in the value of their data.

Does it mean the same thing to everyone?

Even when a governance program is well established, different people may still interpret data in different ways. A marketing analyst may see data as customer behavior patterns, while a database engineer views it as fields and tables. Executives focus on how that same data ties to business outcomes and KPIs. These perspectives are natural, but without shared understanding, they create friction, confusion, and inconsistent results.

Business and technical teams speak different languages when it comes to data. Technical users think in terms of sources and transformations, while business users care about meaning and context. A business glossary bridges that gap. It defines the language of data across the organization, organizes terms, and ensures everyone uses consistent definitions. It also captures synonyms and variations in how terms are applied, helping prevent confusion across reports and systems.

A well-managed glossary connects technical accuracy with business context. It links definitions to ownership, policies, and data quality rules so people can understand not only what the data is, but also how and why it's used. Over time, it should evolve with the business, reflecting changes in meaning and showing dependencies across processes.

Master Data Management (MDM) extends this alignment by synchronizing key domains — like customer, product, or location data — across systems. It turns shared definitions into consistent records everyone can rely on. Together, governance and MDM ensure data meaning stays consistent, no matter where it lives or how it's used.

For organizations refining their data governance programs, this alignment is even more critical. Models and analytics depend on clearly defined, accurate, and synchronized data to perform reliably. When business terms are ambiguous or systems aren't aligned, those inconsistencies flow directly into AI models, leading to bias or unreliable results. A strong combination of governance, glossary, and MDM creates the trusted data foundation that both people and AI systems depend on.

Where does the data come from?

As governance processes mature, it is essential to ensure that data lineage – or data provenance – continues to deliver clear visibility into where data originates and how it changes. Understanding data's journey from its source to its final use is central to maintaining effective data governance. It shows where data came from, how it transformed, and who depends on it – helping teams validate pipelines, troubleshoot issues, and ensure trust as systems evolve.

There are two primary ways to view lineage:

- The **directional view** traces data from its origin to its current state. It shows where data was created, how it moved through systems, and what transformations or merges were applied.
- The **impact view** looks forward to where data goes next and who is affected when it changes. It helps teams understand which reports, systems, or models rely on a dataset — and what happens if that data changes.

Together, these perspectives provide a complete picture of how data flows through the organization. As environments grow more complex, spanning cloud platforms, analytics tools, and AI pipelines, maintaining accurate lineage becomes even more important.

Lineage also serves two key audiences — technical and business:

- **Technical lineage** focuses on data movement within systems, pipelines, and transformations. It helps maintain accuracy and resolve errors.
- **Business lineage** highlights meaning and context: how data supports decisions, which rules govern it, and who owns it. It makes complex data environments understandable and actionable.

Both perspectives are essential. Technical teams need detail to maintain systems, while business teams need clarity to make confident, data-driven decisions.

Lineage is critical to AI transparency and accountability. As organizations adopt AI and machine learning, it's vital to know what data trained a model, how that data was transformed, and what rules apply to its use. Mature lineage provides the traceability that ensures AI models and the insights they produce are explainable, auditable, and trustworthy.

Can I trust it?

Maintaining trust is the foundation of every data-driven decision. Yet most organizations still struggle to answer a simple question: How reliable is our data?

Trust begins with understanding data's condition. Data profiling helps teams assess whether data is complete, in the right format, and usable as it is. Profiling provides the first look at structure, patterns, and anomalies, helping identify issues before data is applied to reporting, analytics, or AI.

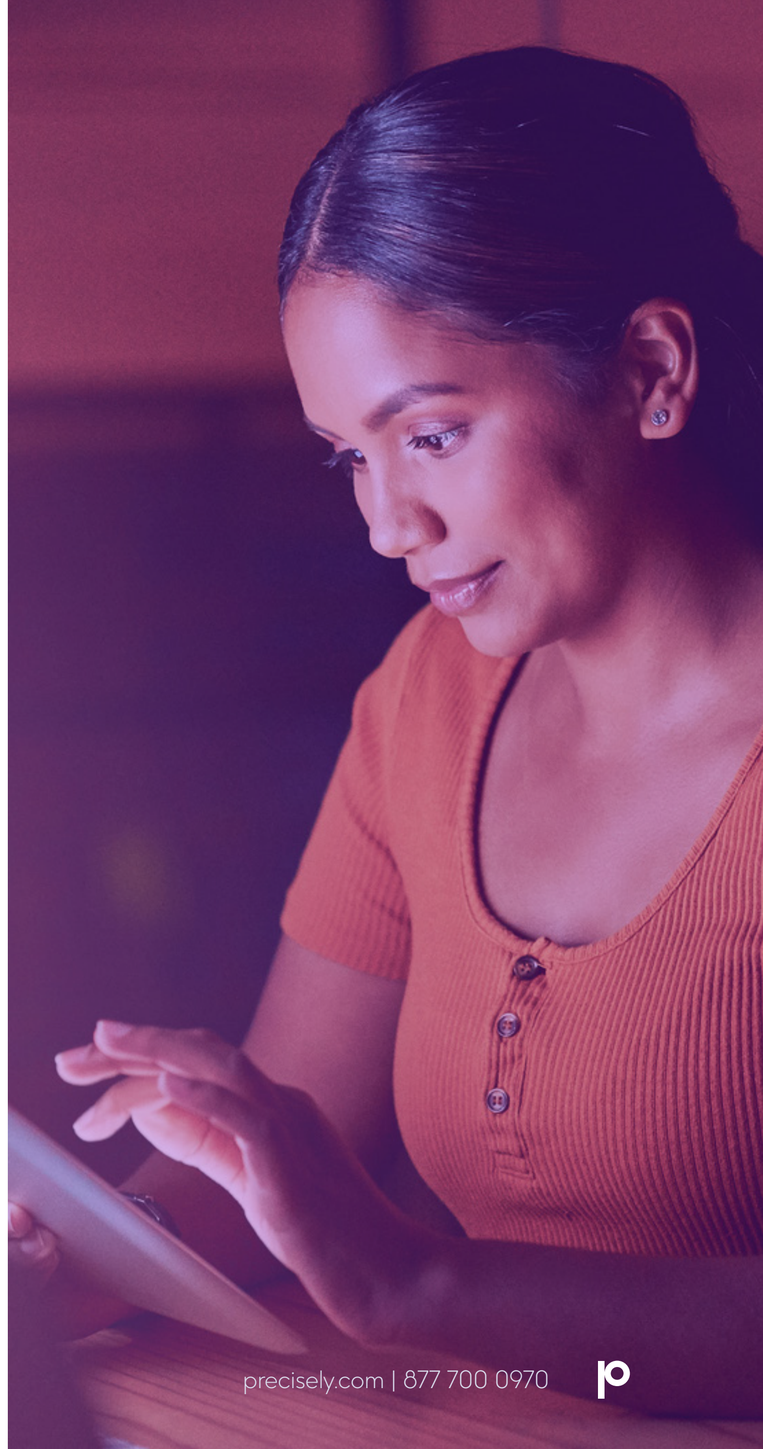
Data quality goes further by ensuring data is accurate, consistent, and free of duplicates. Without strong quality controls, even the best dashboards or AI models produce unreliable insights. High-quality data doesn't happen by accident — it results from clear business rules, validation processes, and defined ownership.

Measuring data quality

To manage quality effectively, organizations need governance that goes beyond detection. Modern approaches calculate quality scores, tie them to business definitions, and alert data owners when thresholds are breached. When issues arise, consumers can reference the glossary and lineage to understand context.

This measurement is especially critical for AI and machine learning. AI models are only as reliable as the data that trains them. Poor-quality data creates biased or inaccurate models, leading to flawed decisions and compliance risks. Continuous quality measurement and monitoring ensure that the data feeding models remains accurate, complete, and current.

Governance provides the structure to make this possible. It connects each quality metric — accuracy, completeness, timeliness, and consistency — to ownership and accountability. When data quality drifts, teams know who is responsible and how to correct it.



Wrapping analytics around data governance

Governance gives organizations visibility into their data; analytics makes that visibility actionable. Together, they move governance from a compliance exercise to an ongoing source of insight and control.

Analytics strengthens governance by automating detection of anomalies, risks, and trends. AI extends these capabilities through predictive governance — anticipating issues, detecting bias, and recommending corrective actions before they impact decisions.

Data observability reinforces this trust by continuously monitoring data health across pipelines and systems. Observability tools provide real-time visibility into data freshness, completeness, and accuracy. When issues occur, teams can detect and resolve them quickly, keeping data — and the insights it drives — reliable.

Regulations such as GDPR and the EU AI Act make this even more important, requiring transparency into where data lives, how it's used, and how AI systems are trained and monitored. Analytics and observability built into governance frameworks expose gaps and reduce compliance risk.

An AI-powered, integrated governance solution brings these capabilities together — profiling, quality measurement, analytics, and observability — to provide a real-time view of trust across your data ecosystem. With these elements working together, organizations gain a comprehensive understanding of their data's health and can use it with confidence.

Can I find it?

No matter how advanced a governance framework becomes, many organizations still struggle to find and understand their data. Knowing data exists is one thing; knowing what it means, where it comes from, and whether it's fit for purpose is another.

That's where data discovery comes in. A data catalog provides a central inventory of your organization's data assets — showing what data exists, where it lives, and how it connects to business processes. It offers critical context such as definitions, ownership, quality scores, and lineage so users can decide whether a dataset meets their needs.

A data marketplace extends discovery by making curated, high-quality datasets easy to find and access. It allows teams to share and reuse trusted data across the business while maintaining governance and control. Together, the catalog and marketplace make discovery faster and more intuitive.

Modern discovery tools make this experience even simpler through natural language search, allowing people to ask questions the way they would in conversation, like "Show me customer churn data for EMEA" or "Where does our revenue forecast data come from?" AI enhances this process by automatically tagging, classifying, and ranking data assets based on relevance, quality, and usage.

Finding data is only half the story. To trust it, users also need to understand its context and purpose. That's why organizing and defining metadata is critical. Governance connects every dataset to its meaning, business definitions, and data quality metrics, ensuring that what users find in a catalog or marketplace is accurate, current, and aligned with business goals.

Who do I ask?

In mature governance environments, accountability becomes the defining factor; who owns the data, and who answers for it? In many organizations, the default answer is still “ask IT,” but IT can’t own the business meaning, quality standards, or policy decisions for every dataset. Without clear accountability, questions go unanswered, projects stall, and trust erodes.

That’s why ownership must remain a foundational principle of effective governance at every stage. Data ownership defines who is accountable for how data is used, managed, and valued across the business.

Every data governance program must rely on three key roles working together:

- **Data Owners** are accountable for the business use of data. They make final decisions about definitions, policies, and access, ensuring data aligns with strategic goals and compliance requirements.
- **Data Stewards** maintain data quality and compliance. They apply business rules, monitor usage, and ensure data remains accurate and consistent across systems.
- **Data Custodians** manage the technical side — storing, moving, and securing data, and maintaining the systems that support it.

Together, these roles form a continuous chain of responsibility. Owners set direction, stewards uphold quality, and custodians ensure security and availability. Governance tools such as glossaries, catalogs, and lineage views make this accountability visible. Instead of guessing who to ask, data consumers can see exactly who owns a dataset, who enforces its rules, and who can answer questions about its meaning or use.

Why Precisely?

Whether your organization is building a data governance program or refining an existing one, Precisely helps transform governance into a connected, AI-ready system that delivers measurable business value.

Precisely is the global leader in data integrity, helping 12,000 customers — including 95 of the Fortune 100 — in more than 100 countries ensure accuracy and consistency in their data. Precisely's data governance solutions create and strengthen business-first frameworks that help organizations find, understand, and manage their most critical data with greater transparency and efficiency.

The AI-powered Data Governance service within the Precisely Data Integrity Suite provides a business-friendly, SaaS-based framework to build or enhance your governance initiatives. It enables teams to discover, trust, and leverage critical data across the enterprise with quick implementation and intuitive configuration. It features a flexible metamodel that accelerates adoption and adapts to diverse use cases.

Through the Data Integrity Foundation, the service connects seamlessly with other Data Integrity Suite services — such as Data Quality and Data Observability services — to deliver a unified view of data health and reliability. Automated metadata harvesting and AI/ML-driven tagging, classification, and relationship mapping, boosts productivity and collaboration across teams.

With Precisely, organizations at any stage of their data governance journey can build confidence, strengthen accountability, and unlock the full potential of their data. The result is an integrated, scalable path to data integrity — empowering you to minimize risk, improve insight, and give your AI initiatives the trusted data they need to succeed.

Contact us to optimize your data governance program today.

precisely

About Precisely

As a global leader in data integrity, Precisely ensures that your data is accurate, consistent, and contextual. Our portfolio, including the Precisely Data Integrity Suite, helps integrate your data, improve data quality, govern data usage, geocode and analyze location data, and enrich it with complementary datasets for confident business decisions. Over 12,000 organizations in more than 100 countries, including 95 of the Fortune 100, trust Precisely software, data, and strategy services to power AI, automation, and analytics initiatives. Learn more at www.precisely.com.

www.precisely.com