

Generative AI and data governance: a powerful combination for business success

Forrester Analyst, Raluca Alexandru, joined Collibra's podcast, *The Data Download*, to answer questions on AI and data governance.



How do you get started with a data governance program?

Getting started with a data governance program involves several practical steps to ensure its successful implementation. The specific steps and priorities may vary depending on the organization's size, industry, and maturity level; however, certain elements can provide structure to the process of designing and implementing your program.

As with the start of any journey, setting clear and specific goals is key. For the overall governance vision, it's important that you have a clear understanding of why you are embarking on this journey and where you want to go in the medium to short term. Once your goals are clear, make sure that you understand what they apply to — map out your data universe, critical assets, business processes, and stakeholder interactions to ensure that your goals are comprehensive.

Once your goals and mapping are in place, start designing the data governance team structure. Two things are important here. First, based on your stakeholder and data assets mapping, you are aiming at a cross-functional team that is representative of your internal stakeholders. Think of roles spanning IT, business, compliance/legal, data management, and architecture. Second, establish a governance leadership role that can bring the right experience and qualities to efficiently leverage the relationship with not only your stakeholders and the leadership/executive team but also the extended teams. This will ensure buy-in at the executive level and also increase adoption and platform usage — and, therefore, performance — within the company. The roles needed in your data governance team or council often already exist within your company. Assess the current roles in your organization that are involved in data governance responsibilities and run a gap analysis of the missing resources. The team will develop along with your efforts. Forrester offers a detailed framework for connected intelligence teams, with clear roles and responsibilities. Reach out to us for more information.



How do you get started with a data governance program? (continued)

Once the team is in place, start mapping the organization's current landscape. A thorough data assessment, including data sources, quality rules, accessibility, and security policies and risks, will ensure that the team is ready to identify the organization's current governance structure and act on it to take it to the next level. It is a common misconception that unless an organization has a clearly established data governance program, there is no governance happening in its data landscape. The lack of a framework also translates into a data governance model — though not a very performant one.

There are several data governance models and frameworks out there that can serve as inspiration. Go back to your goals and the specifics of your data landscape and decide on a framework that works for your organization. Some organizations with large legacy data sets and processes and legacy technical platforms can find it daunting to reorganize their data assets and flows. Have a critical look at what you can reuse and where to start fresh. While it's not an easy decision to make, it can save the success of your program and the resources that you will need to invest in it. Emerging processes can run in parallel for certain periods of time, and new policies can be tested against different scenarios. Think of data policies and standards that will guide your data-handling practices, privacy, access management control, security and risk, and compliance.

As part of your data processes design or revamp, you will need to focus on elements such as data classification, data quality management, metadata management, data access management, and overall data lifecycle views. Next to that, you will need to think of how you can communicate and implement these processes for different stakeholders in the organization. Having a clear understanding of the roles and their responsibilities will help you map out data ownership and access and approval rules as part of the stewardship program.

Keeping your goals and stakeholder interests at the core of your program is essential. While you need to get traction with the leadership team, it's also important to increase adoption and usage of the platform from the business and even legal and finance (think environmental, social, and governance [ESG]-related governance use cases). Communicate on a regular basis about the goals and results of your program, translate technical achievements into business terms, and set up a companywide data literacy program that allows your stakeholders to understand your framework and get involved.

On the platform side, you might already have a solution or be looking to acquire one. This is a conversation for next time. Let's focus on setting up your data catalog and metadata inventory. The data catalog, while not enough on its own, acts as the engine of your data governance program. You want to make sure that you run a comprehensive data discovery process that imports or maps all your data assets, business definitions, technical metadata, ownership, and lineage flows. Next to this, you want to have a clear overview of your data systems, data repositories, applications (internal and external), and tools that you currently use in your organization. This is important when designing access control management measures and identifying and mitigating security risks.



How do you get started with a data governance program? (continued)

Once your data catalog, technical metadata, and business glossaries are in place, you need to define your stewardship roles. Assign clear responsibilities for data stewards and decide on ownership of data domains. Data stewards will be the owners and guardians of the data domains, ensuring that quality, integrity, and access are under control. For support on this, check the RACI matrix that Forrester offers for different stewardship roles.

When it comes to specific processes such as data security, privacy, and quality, start by aligning your overall business goals to the type of data assets you are managing. Then, set the relevant metrics to help you monitor and improve your processes — think about data quality metrics, data cleansing and remediation strategies, and data profiling. For security, make sure you have correctly mapped out your sensitive data, defined the policies that comply with relevant regulation, and put measures in places that can help you mitigate data breaches — think of access controls, data encryption, and monitoring mechanisms here.

To round things out, look at your data governance program as a circular effort. As part of your governance cycle, you want to ensure continuous and consistent training programs as part of your data literacy to keep your stakeholders knowledgeable and engaged. At the same time, you want to be able to continuously monitor and measure the success of your program and platform adoption. The combination of the two will provide you with valuable insights that you can feed back into the framework design process to optimize results.



How do you measure the value of a data governance program?

Measuring the success or value of a data governance program depends mainly on the set objectives. Clear objectives will lead to specific metrics and measurement practices that, in turn, will deliver value and success indicators. You can measure success and adoption and the corresponding metrics in areas such as compliance, where we usually look at things such as relevant local/industry-specific regulations; security standards and policies; and data privacy practices and regulations. For data quality metrics such as data consistency, you'll find that accuracy, completeness, and validity are often relevant. In the context of security and compliance, data access and usage are elements that are often measured and used for mitigating risks — data accessibility and consent management policies are popular metrics. The overall purpose of setting metrics and measuring against them is to observe what is happening with your data in real time. You also ensure a continuous feedback loop of the data and usage insights back into your program, leading to performance optimization, an increase in value, and long-term adoption of governance efforts.



How can a good data governance program help companies safely, legally, and ethically take advantage of the potential of this rapid rise in generative AI/large language models (LLMs) to return the most value?

Generative AI and the governance around it is happening as we speak. With quite a few changes in regulation expected, the market is following client demand and technical innovation speed. Following the principle of AI governance, the focus is still on controls and clear processes and policies around data discovery, profiling, data quality, security and compliance policies, stewardship, and collaboration. The quality, security, and reliability of the data input for AI models, including generative AI functionalities, dictate the level of success, safety, and legal and ethical compliance of the applications.



How do you connect your catalog and business glossary to company business goals (to drive value)?

Connecting the data catalog and business glossary to business goals is key in ensuring the success of a data governance program. The first step is making sure that you have clearly defined business goals and they're in line with the overall company strategy. Understand the goals and how they translate in governance terms. Define metrics for each business goal and assign data categories that contribute to achieving them. During data discovery and cataloging, clearly map and label data and technical metadata to match company business goals. The business glossary will closely follow the discovery and mapping process to provide clear and comprehensive definitions of how the business understands the value of the data assets. Ensuring alignment between business expectations and the catalog structure is essential for the long-term success and adoption of the governance framework.



If you are interested in learning more, please visit our [podcast webpage](#) and listen to the Forrester episode