

It's Time to Rethink How We Govern Data

A Step-by-Step Guide to Modern Data Governance



Copyright © 2019 by Pacific Data Integrators

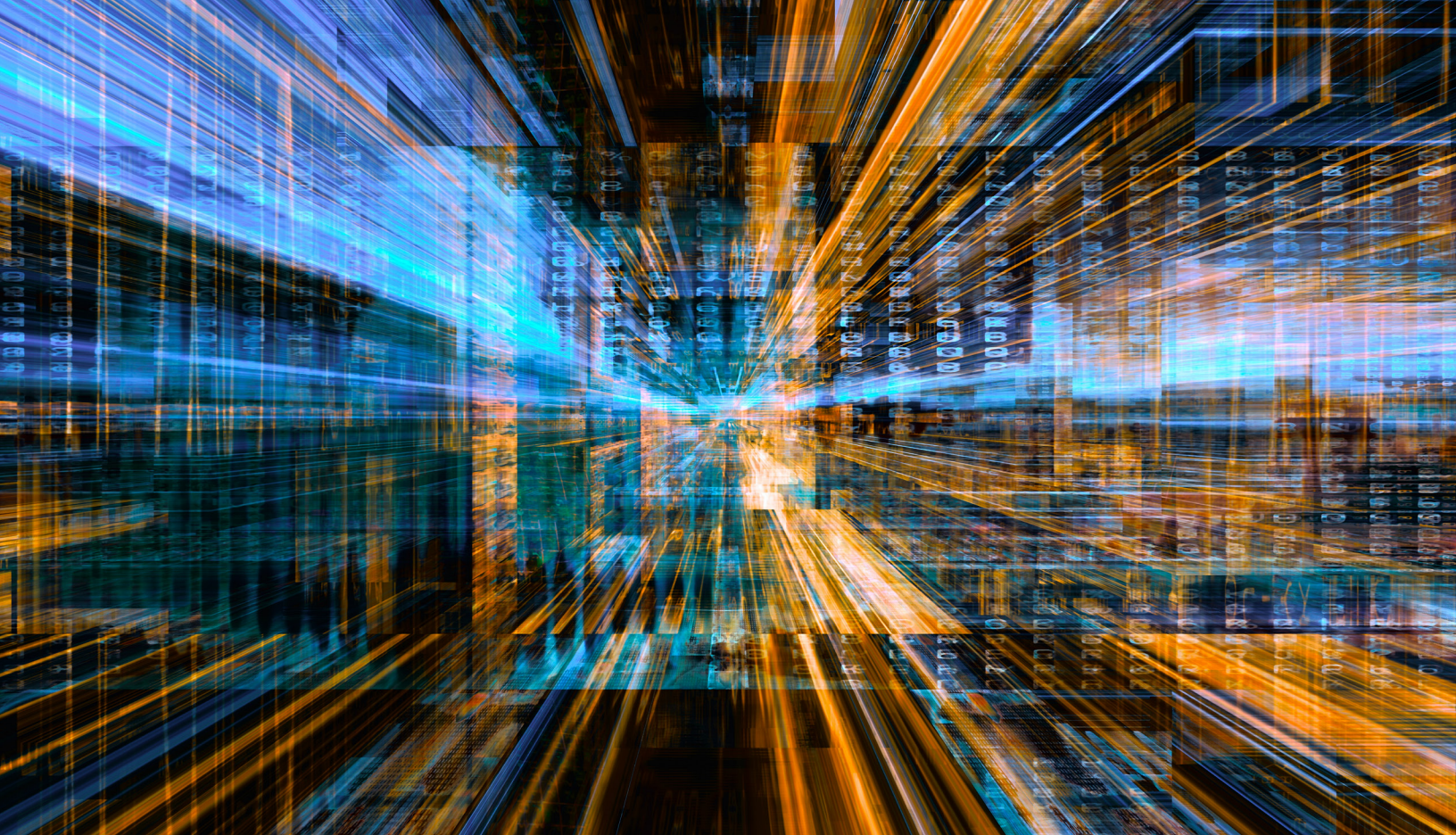
All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law. For permission requests, contact Pacific Data Integrators at the address below.

Pacific Data Integrators
5710 Lone Tree Blvd
Rocklin CA 95765
pacificdataintegrators.com



Table of Contents

Modern Data Governance	3
The Cost of Poor Data Management	3
A Better Approach.	4
Scaling Data Governance	5
Who Data Governance Benefits	6
8 Steps to Implementing Data Governance	7
Step 1. Define Data Governance Framework	7
Step 2. Conducting a Maturity Assessment	8
Step 3. Choose & Implement Data Governance Tools	9
Step 4. Prioritize Business Problems	11
Step 5. Contextualize Business and Technical Data Assets	11
Step 6. Implement Data Quality.	12
Step 7. Publish a Data Catalog	12
Step 8. Develop Program Metrics and Reports.	13



Modern Data Governance

The Cost of Poor Data Management

Producing, acquiring, and/or publishing data is a huge cost to organizations. Companies spend millions of dollars on the people, processes, and technologies involved in managing data. It's an extremely valuable investment—but only when the resulting data is used to guide business decisions!

Unfortunately, many businesses spend money on data management and then fail to use that data in any meaningful way. If company information is not being used for decision-making, then why are so many organizations spending so much money storing and managing that information?

Even when companies try to use their data, there is often friction between what data consumers require and what they receive. From the report owner who prepares the data to the marketing team that examines the data to discover

customer trends, these people profit greatly from clean, well-managed data. Projects are completed faster and business decisions are more effective. So, what gets in the way of these people and using their company's data?

The issues that create friction for data consumers are straightforward to identify, if not to solve. Data cannot be used if it cannot be found, is of poor quality, keeps changing, cannot be trusted, or if the context surrounding the data is missing. Often data is duplicated and stored in many different environments, creating too many sources of the "truth."

At the base of this is visibility and trust. Employees cannot use data if they cannot access it, and they will not use data if they do not trust it.

How do we fix this?

A Better Approach

Data governance allows organizations to unlock the value of their data. An effective data governance program ensures the right people can access the right data and—most importantly—it builds trust in that data.

Several prominent data trends have contributed to a revitalization of data governance. First, companies across all sectors have recognized that data is a strategic asset. Companies that treat data as a valued resource, and invest in leveraging their data, have tapped into digital transformation. This isn't just another buzz word – it's a strategy that the largest, most innovative, and most ambitious companies are using to disrupt their industries.

At the same time, the sheer volume of data is growing. Businesses today collect more asset, customer, and product data than ever before. The problem today is not typically a lack of data; it is taking the deluge of information and creating something useable out of it.

But as companies try to leverage their data, an increase of data-centric regulations seems to be working against them. At least, that is how it is perceived. Companies attempt to gather as much information as possible, while governments push back with limitations and regulations.

Such regulations do not have to be viewed as an extreme hindrance, or even a hindrance at all. Savvy companies are using these regulations as a launch pad for modernizing their data governance programs. Modern data governance programs are used to fortify data security policies, rebuild customer trust, and more.

Before, everyone tracked data assets in Excel or SharePoint. This method is tedious to maintain, as it takes weeks of manhours to document. Even worse, static data governance documents are out of date the moment they are published. Tracking data governance this way was never effective, and it has become even less effective now that companies are on the Cloud, are gathering big data, and are creating hybrid environments.

Data is complicated. Tracking its lifecycle on an Excel sheet is like writing code on a piece of paper. Is it possible? Sure. But it's wildly antiquated and at the end of the day, not very useful.

Today, data governance tools promote collaboration, enabling all team members to view the flow of data, its stakeholders, quality, and more. They move quickly, adjusting to new regulations and generating compliance reports in minutes rather than weeks.

Modern data governance allows data to be identified and detected throughout its lifecycle, empowering data security programs to safeguard information from its creation to archival. And—perhaps most importantly—modern data governance tools are built to scale.

Tracking data's lifecycle on an Excel sheet is like writing code on a piece of paper. Is it possible? Sure. But it's wildly antiquated."

Scaling Data Governance

When looking to change, large corporations are hindered by existing technology, established processes, and a plague of silos that stop collaboration in its tracks. These massive, entrenched business challenges make implementing a data governance program seem overwhelming. Like a pipe dream, data governance would be wonderful to have but appears impossible to implement.

These challenges are why modular—and therefore scalable—data governance tools are a game changer. Scalability empowers even the largest companies to start small and implement data governance one project at a time. This allows stakeholders to prove the ROI of data governance and get the backing of everyone in the company. No one wants a new program or tool thrown at them, a program that becomes costly to manage, or a tool that doesn't get used and sits around gathering dust. By starting small and growing your data governance platform on a healthy repertoire of successful projects, you can generate a culture where data governance is valued and utilized.

Who Data Governance Benefits

Modern data governance benefits employees across the enterprise. The most obvious benefits are to legal and governance officers, who are looking to protect the company from fines and legal suits by ensuring compliance. New regulations can be ambiguous and open to interpretation, meaning data security policies may change several times as existing regulations are clarified and new legislation is passed. Without agile data governance, altering data policies is a cumbersome process.

However, modern data governance affects many other departments as well. It benefits line of business individuals, who need to ensure their analytics can be trusted. This means knowing where their data came from and the quality of the data.

Then there are data stewards, who need to manage metadata and would benefit from a method of tracking data quality through the lifecycle of the data. Without data governance, it is difficult to integrate data across cloud and hybrid environments and get a clear picture of a company's information flow. With data governance, a business can reduce maintenance

costs while delivering connected and trusted data to the individuals who need it.

90 percent of executives say data governance is important to big data business adoption.

Agile and robust data governance benefits data analysts and architects as well, as they need to discover and understand the data they are using. Gone are the days

where data governance restricted employees and was comprised of top-down policies. Today's data governance empowers employees, if they are taught to leverage and value it. It is democratized and built to increase the usage of value data.

This is why 92 percent of executives say data governance is important to big data business adoption. When data governance is responsive, collaborative, and transparent, it supports employees across the entire enterprise.

Parallel to this growth in data governance is increasing recognition of the Chief Data Officer (CDO) role. More companies are creating this job position, and those that already have CDOs are tasking them with more responsibilities and more ambitious projects. Not only is this a boon for data governance, it is also creating a huge career opportunity for information management professionals who understand the importance of data.



8 Steps to Implementing Data Governance

Step 1. Define Data Governance Framework

A healthy enterprise data governance program starts with a strong understanding of what “enterprise” and “data governance” means for the organization.

Start by asking questions to understand which business units and organizations will be subject to data governance. Look at the big picture—when a truly “enterprise” data governance program has been implemented, what will that cover? Where are the existing data-centric skills? What is the current view of data ownership? What is the company culture?

After outlining the initial data governance framework, the team needs to decide the data priorities. Is there a regulation the company needs to become compliant with? Is there a certain department that is struggling to visualize their data? There are numerous departments and countless projects that can be

used to launch a scalable data governance program. Choosing a specific business case to start with is an important determination that will guide the team in launching the data governance initiative (see *Scaling Data Governance*, above).

Defining the framework will provide direction, but it is also important to define what “data governance” means to the company. There are many aspects of data governance, some of which will be more important than others depending on the goals of the company. Curating the definition of data governance into something that resonates with a company will not only guide the program towards solving the business’s challenges, it will also cultivate support.

After defining the framework, the team needs to identify an executive sponsor. Sponsors are crucial leaders that provide credibility, rally support, and drive projects forward. An executive sponsor will work to prevent data governance from stalling and will be a beacon for growing a culture of data governance.

The last step to this process includes identifying stewards. Each data steward owns and manages a discrete part of the framework. The data governance process will involve data stewards, but ultimately data stewards will be directly accountable for the success of the management of their data domains.

Step 2. Conducting a Maturity Assessment

Defining data governance provides a structure and direction for data governance. Conducting a maturity assessment builds a foundation upon which to build.

- Crucial aspects to assess include:
- The data management maturity for the organization
- The capacity for change
- The ability to collaborate
- How business uses data within its strategy

This assessment will show an organization its strengths and weaknesses and will pinpoint the largest business challenges. If there are weaknesses in one area, knowing this from the start will allow the data governance team to implement additional measures to secure the success of the data governance program.

This is also an opportunity to develop a use case backlog, which will be important in step three.

Step 3. Choose & Implement Data Governance Tools

Once a data governance framework has been defined, a company's next step is deciding on the data governance tools that will be used to implement and maintain the data governance program. This is a vital step, as ineffective tools will cripple even the most robust data governance plan.

Informatica Data Quality

Most data experts agree that data quality is the best place to start for companies that want to digitally transform how they govern data.

It is difficult to govern inaccurate data, and even when it is governed, business analysts will not trust the data enough to leverage it. Having high quality data makes the effort of data governance worth the investment because the data will be trusted by analysts, and the resulting business decisions will be trusted by executives. It is also difficult to secure a company's sensitive information when the business is unsure where all of its sensitive information resides.

It is for these reasons that data quality is an important factor in creating a data governance program. In fact, when data quality and governance are integrated, data quality helps data governance run more smoothly and more efficiently. For example, while Informatica Data Quality works to match and deduplicate data, resolving errors in company databases, the data governance tool AXON allows data stakeholders to see the data quality as information moves from system to system. Companies can see where there are drops in data quality and determine where data quality needs to be improved.

Enterprise Data Catalog (EDC)

You cannot govern or protect data if you do not know that the data exists or where it is held. Discovering a company's data can be an enormous challenge, due to the rise of Cloud, big data, hybrid environments, and more. Even before this explosion of information, data siloes were enough to keep information hidden.

The solution to this is an enterprise data catalog. Informatica's Enterprise Data Catalog (EDC) is a powerful tool that scans a company's data to determine what data exists,

where it resides, its quality, the processes that it touches, the stakeholders in charge of the information, where the data originated, and more.

Data stewards may think they can uncover these answers manually, and perhaps once upon a time they could. However, the sheer volume of information that flows through a modern enterprise demands a tool that can catalog data at a large scale.

Manual efforts cannot discover, map, or track an entire enterprise's information flow. EDC leverages the power of artificial intelligence and machine learning, providing visibility into all data, wherever it resides.

This is why EDC is an essential tool for any data governance program. Informatica has taken a monumental—and error prone—manual effort of discovering data and replaced it with a robust, intelligent tool. Only after EDC shows a company their data can they begin to govern it.

AXON

Data governance is a framework used in the management of an enterprise's data assets. Data governance holds data to a certain standard, ensures key stakeholders have visibility into the quality of their data sets, and is essential to implementing and reporting on data privacy regulations.

Data governance tools are becoming increasingly important assets to businesses, considering how vital data is to maintaining a thriving enterprise.

Historically, companies have had no choice but to manage data governance on spreadsheets or SharePoint. However, this method gives no insights into the quality of managed data, is tedious and time consuming, and is constantly out of date.

Companies that invest in data governance platforms reap many benefits of well-managed data, from compliance to more accurate business decisions. There is a reason that the AXON data governance platform is Informatica's fastest growing product in the history of their company. Simply put, data governance has a proven ROI.

Since AXON is a scalable data governance tool, companies can start small and then scale their governance programs. After one project is used to prove the ROI of data governance, data governance can be expanded to other projects and departments and gradually expand its reach until the company has true enterprise data governance.

Step 4. Prioritize Business Problems

This is the time to identify critical reports, KPIs, prioritized data sources, etc. Identifying the goals and issues of a company will allow the team to identify specific benefits (especially financial benefits) that can be achieved through data governance. The overall goal of this step is to show how data governance can improve the usefulness of data and content.

By identifying business use cases (along with the appropriate steward), the team can build support for the project, to gain momentum, and to lay the groundwork for a culture of data governance. By finding the areas that need the most help and tackling those first, the road forward will be smoother thanks to provable ROI and a visible increase in efficiency.

Begin by interviewing users directly; this is the best way to understand their business challenges and define what high quality data means to them. This high-level assessment is often conducted in the form of surveys or interviews. Next, evaluate the architecture and systems that surround those business challenges, in order to outline a prioritized list of challenges. This prioritized list can be organized based on critical data projects, regulations, critical data systems, or known data issues.

The goal of this step is to dive deeper into the reasons that end users are having data difficulties. For instance, an end user might report that it takes two weeks to get their data. In that case, the team would examine the architecture design to better understand why it's taking so long.

Step 5. Contextualize Business and Technical Data Assets

Once the team has their prioritized list of use cases—with assigned stewards—from step 3, it is time to flesh out solutions to the business challenges. With the context provided by end users, the data governance can now bring context to the data and craft solutions to overcome the reported challenges.

Bringing context to data simply means identifying which data elements are critical, and assigning a criticality index to further refine the importance of the element. Additionally, creating business terms based on the CDEs and implementing a glossary for the prioritized business cases will allow stewards to govern the assets.

But contextualizing data assets doesn't stop at creating and defining business terms for the CDEs. The data governance team needs to work with the stewards in order to document

end-to-end data lineages and also to associate business terms and system metadata to business intelligence analytics and reports. Tying these relationships together in Axon and EDC unlocks the full potential of data, giving users the ability to truly understand how data is used in the organization.

Step 6. Implement Data Quality

Data experts agree that strong data quality is essential to creating effective data governance. The problem with tackling data governance before data quality is that employees will not trust the data they are using. The data may be governed, but that does not mean it is accurate.

Essentially, there is little point in creating reports or the like if employees do not trust the information enough to actually make decisions based on what their analytics say. This is why most data experts agree that the best strategy is to tackle data quality before launching a data governance program. This is widely considered the best way to create a healthy culture of data governance. So before moving forward with a data catalog, consider implementing data quality, if it is not already in place at your company.

Step 7. Publish a Data Catalog

By this point, a company will have a strong understanding of the business challenges end users are facing and the context surrounding that data, and they will have a strong foundation of quality data to rest their data governance program on. At this point, it is time to publish the governed data in a catalog, like EDC, for other consumers.

Publishing the data in a catalog democratizes the data, allowing the end users who leverage it to view the data, its quality, where it came from, and more. Instead of grappling with multiple sources of the truth, the data governance team can certify data as accurate, building trust with the end users leveraging the information.

For a data catalog to be most effective, the data governance team will need to deliver education to all consumers, empowering them to find, use, and trust the assets in the data catalog. This unlocks the value of the corporation's data and enables better decision making.

Step 8. Develop Program Metrics and Reports

The final step in this data governance model is vital to reporting the progress and value of the data governance initiative. The data governance team wants to communicate the value of its initiatives to the entire enterprise in order to build a pro-data governance culture. In order to specifically show how data governance added business value and met its stated objectives, the team will need to gather metrics including:

1. How has data governance solved end users' reported business challenges?
2. How has data governance reduced risks?
3. Has the program improved operational efficiency?
4. How often are stewards using the relevant tools?
5. Have the training and educational efforts been effective?
6. What are user adoption rates? If necessary, how can these rates be improved?
7. Are the implemented policies and processes working appropriately?
8. Is the staff following the provided data catalog guidance, and changing their behaviors when managing data?

These metrics will help the data governance team communicate their success to the rest of the company. It cannot be stressed enough that data governance success must be shared. If departments have visibility into improved efficiency and business decisions made by others leveraging data governance and a data catalog, they are much more likely to get on board the data governance program and make the effort to learn how to leverage a data catalog. On the other hand, if they've never heard of data governance and have no idea how it could improve their reports and processes, why would they put their full energy into implementing it?

Communicate, communicate, communicate.

These metrics will also aid the data governance team in optimizing their program moving forward. It's important to revise the documented data governance strategy and roadmap as the program matures. Plan for roadmap changes and repeat steps 4–8 when implementing each use case.