# Why a data steward is crucial for effective data governance

- Data stewardship vs. data governance
- Why is a data steward needed?
- How does distributed data stewardship work?
- Keys for successful implementation

TechTarget

# Data stewardship: Essential to data governance strategies

*GEORGE LAWTON, CONTRIBUTOR*

Organizations are establishing data governance programs to address evolving privacy requirements and increase the utility of business data. The lack of a data steward can undermine any data governance effort.

Approaches to data governance vary depending on the organization. Some rely on tools to keep data in line. Others centralize data governance under roles like CIO. Others ask data product managers to drive the value-creation aspects of data governance.

Data governance incorporates controls to help understand metadata, data quality and proper use, said Matt McGivern, a managing director at Protiviti, a management consulting firm.

"It has become popular to call data an asset, but without proper data governance, this is impossible," he said.

Data governance is a strategic analysis technique, while data stewardship is a more tactical role. Data steward could be a specific job title or a responsibility within the job

of a business manager. Every organization should lay out its data governance goals and specific measures to translate them into daily practices.

To be an asset, data needs to be understood by an organization. The organization must know where it's stored, how it's utilized, who is utilizing it and when it is destroyed or disposed of. Data governance supports each of these areas, and data stewards drive understanding of appropriate data use. They also establish the standard controls to data and ensure that it is used in the appropriate context by confirming contractual restrictions with legal guidance.

Depending on the situation, the differences between data governance and data stewards determine if one is needed. Organizations also should consider what may happen if they don't have a steward in place.

## DATA STEWARDSHIP VS. DATA GOVERNANCE

Every company has its own distinctions between data stewardship and data governance. Most experts agree data governance is a broader concept and data stewardship is a specific role to put it into practice. There are different ways of framing the distinctions between them.

**Broader versus supporting controls.** Data governance is a comprehensive set of controls -- strategies supported by policies, procedures and technologies -- to help control data, McGivern said. Data stewardship, by design, is a specific supporting set of controls to help data governance act.

"Data stewards are often where 'the rubber meets the road' for data governance controls," McGivern said. "[They are] helping to provide the necessary context to the data and other knowledge about required controls, proper usage and the current state of quality."

**Framework versus role.** Data governance establishes a framework for how an enterprise provisions and stores data, while data stewardship is a role within the organization that advocates for effective uses of that data to create value, said Ed Murphy, senior vice president of data science at 1010data.

In this regard, governance focuses on the essential infrastructure for data-driven enterprises:

- Roles.
- Access controls.
- Storage.
- Availability.
- Metadata integration.
- Disaster recovery.

- Legal and regulatory compliance.

In contrast, stewardship focuses on wringing value from data:

- Data set discoverability.
- Documentation and metadata.
- Change management.
- Quality advocacy.

**Policies versus tasks.** Data governance is the ownership and management of the policies, processes and procedures for data, said Claire Thompson, chief data and analytics officer at Legal & General Group, a financial services company. In contrast, data stewardship is undertaking tasks needed to ensure adherence to and compliance with the data policies and frameworks. The data steward interprets and implements required processes on a day-to-day basis.

**Outcomes versus inputs.** Data governance is the outcome while data stewardship is the input required to achieve it, said Graeme Thompson, CIO at Informatica. Data governance includes what someone is trying to achieve and the mindset necessary to achieve it. Data stewardship involves defining ownership and responsibility for different pieces of data and for the processes that generate the data.

While they're different, they are not separate.

"Most organizations that are committed to data governance invest in a data stewardship practice," Graeme Thompson said.

This is often handled by a dedicated team, or at least a team that has the responsibility, established processes and supporting technology to manage the data.

**Procedures versus standards.** Data governance establishes procedures while data stewardship ensures reliable standards, Graeme Thompson said. Procedures include classifying data, tracking its lineage for transparency, and enforcing rules for appropriate data access, usage and retention. The standards protect critical data and ensure compliance with appropriate conditions and objectives.

## WHY IS A DATA STEWARD NEEDED?

In theory, an organization can get by with a data governance process spelled out by the CIO or chief data officer. However, experts see several benefits of a data steward over and above a more general data governance practice.

**Apply subject matter expertise**. Data governance applies to the broader field, while data stewards can use it in a specific business domain.

"Data stewardship roles are usually undertaken by people within a specific business area so they would have the SME [subject matter expert] knowledge of what that area does and what systems and data it uses to undertake the key business processes," Claire Thompson said.

Data stewards work with business leads and data owners to ensure regulatory, risk and policy compliance. They must identify, define and document the key data points, including the flow and usage of data. Data stewards can also monitor and analyze data to check quality, identify problems and escalate potential data risks.

**In this handbook:**



**Manage assets that manage the data**. Data stewards protect and secure data because they manage the assets that manage the data, said Naveen Kamat, vice president and CTO of data and AI services at Kyndryl, an IT infrastructure services company. They can also provide insight into which assets are working and which are not. Data stewards can make recommendations on assets to remove or utilize more within the organization.

**"All analytics-driven organizations will be slower without a steward as they try to reverse-engineer a data set for lack of documentation or quality."**

**Ed Murphy**

**Senior vice president of data science at 1010data**

**Facilitate between IT and business users**. Data stewards are sometimes referred to as the data ambassadors within an enterprise because they are the lynchpin between the technical IT operations, data management teams and line of business users who consume business-critical data, Graeme Thompson said. When data is safe, secure and appropriately managed throughout its lifecycle, workers can trust it to make strategic, data-driven business decisions.

**Create value with data**. Data stewards make it easier for companies to treat data as a product, said Bret Greenstein, data and analytics partner at PwC. Data product ownership is all about creating value with data. The data product owner can be the

data steward, adding to their responsibilities. The data product owner focuses on the use and value created from data.

Product success is typically measured by the value of the product, its usage and user satisfaction. Data product owners focus on driving data requirements that make a difference for the business. This helps companies achieve more value from their data.

## PAINS OF NO OVERSIGHT

Not having a data steward can hurt the organization in the following ways:

**Inconsistent naming.** Data stewards can ensure consistency. Greenstein recently encountered a company struggling with customer data. They had it stored in multiple places that did not follow consistent naming. The same customer showed up under different customer IDs for various functions in the business. Customers got different pricing and discounts depending on the channel. The business leaders did not have a holistic view of a customer's buying history, preferences and needs.

**Reduced speed.** Data teams lose speed without stewardship because low-quality data requires more cleaning, more dev time, more code and more testing than high-quality data. Data issues slow the entire development cycle.

"All analytics-driven organizations will be slower without a steward as they try to reverse-engineer a data set for lack of documentation or quality," Murphy said.

**Poor data quality.** Lack of data stewards can lead to poor data quality that inhibits adoption of new technologies and infrastructure, Claire Thompson said. Without data stewardship, organizations can lack effective screening for financial crime obligations and decision-making around business and customer initiatives. The result is inaccurate reporting and inefficient service delivery and outcome for customers.

**Lost opportunities.** Not having a data steward reduces proactive uses for data when companies are trying to enable new business outcomes, Graeme Thompson said. For example, if a customer experience program fails to understand consumer buying preferences across multiple channels, it will miss opportunities to improve products or recommend similar offerings. The lack of a data steward can also make it harder for enterprise data consumers to find data that they can trust.

NEXT ARTICLE

# Enhance data governance with distributed data stewardship

*SUE TRIPATHI*

In this data-driven era, organizations are realizing the significance of governing their data. Poor governance can lead to low data quality and lack of visibility.

Deploying a data stewardship or distributed data stewardship model aids governance, fueling positive business outcomes.

With large amounts of data -- data storage on cloud and otherwise, complex data pipelines and ecosystems, higher scrutiny and more consumption of data -- organizations are moving toward the democratization of data. This enables everyone in an organization to work with and feel confident talking about data. They use observability tools to monitor, track and improve the quality of data in real time. Data stewardship models, armed with customer experience information, have become a hub of data-information decision-making.

Data decision-making depends largely on the use and consumption of relevant data across an enterprise. This often leads to conversations about data ownership, accountability, visibility, quality and the like. It also raises questions about what constitutes a data stewardship model and creates confusion around terminology.

The *data stewardship* and *distributed data stewardship* model lingo is frequently mixed together, with rudimentary understanding between the two. This article aims to dispel any misunderstandings, clarifying how these two terms are different and why organizations should consider using one of these models for their data operations.

## DATA STEWARDSHIP VS. DISTRIBUTED DATA STEWARDSHIP

Data stewards are a central point of contact. They enforce accountability of the data lifecycle, and oversee data governance and visibility. In many instances, data stewardship is a centralized business or IT function. These settings require enterprise data governance or expertise in data management and governance execution.

Distributed data stewardship is a model or framework that allows teams closest to the data to manage access and permissions. Data management is decentralized and resides within the business unit.

## TO BE OR NOT TO BE A DATA STEWARD ORGANIZATION

The promise of data stewardship is usually not fulfilled because many organizations don't clearly define the role of data stewards. Formalizing the role of stewardship is essential to preempt confusion or disillusionment.

Effective data stewardship programs begin with an organization's data management and governance. Key components include the trifecta of data stewardship: people, technology and processes.

- People means the data-centric skills of staff, data analysts and modelers.
- Technology encompasses knowledge of data systems.
- Processes covers comprehension of where relevant data resides and what is available for different projects.

**The success of such a model depends on how well a decentralized IT, governance and distributed access management model works.**

The core component of a distributed data stewardship program is similar to a data stewardship one. The success of such a model depends on how well a decentralized IT, governance and distributed access management model works.

Because a distributed data stewardship model delegates data management responsibilities throughout the enterprise, the fundamental difference between a data stewardship model and a distributed data stewardship model is in shifting an organization toward decentralizing data access. This requires time, effort, cadence and key stakeholders who agree and adhere to such a framework.

In either case, data-driven organizations adopt a customer-focused, measurement-driven, quality or lean six sigma data improvement initiative for their data stewardship model or program.

## HOW DOES DATA STEWARDSHIP AND DISTRIBUTED DATA STEWARDSHIP WORK?

To implement a successful data stewardship program, three core elements must be in place: organizational culture, data management and measurement.

1. **Organizational culture.** Data stewards can only be successful if enterprise executives support and help incorporate their vision. Clarity on data ownership, breaking down internal silos, reusing data and sharing it across the enterprise gives data stewards staying power to realize successful enterprise outcomes.
2. **Data management.** Defining, managing, tracking and improving data quality generate attention on data origin, its architecture, how it's administered, and its security and access rights.
3. **Measurement.** Data stewards must align their work with clear success metrics. These metrics may include financial, quality and human development; labor productivity; and alignment with business objectives. Metrics to track progress in an organization include reduction in complexity, lower cost and execution time, higher performance metrics, optimized value from data assets, improved speed of execution and elimination of redundancy, increased collaboration and improved results.

Without these core elements, data stewardship models may not yield an organization's executive vision about the value of data. This includes what data is relevant, utilized, consumed, reusable and scalable across the organization.

Various real-world examples of data stewardship models indicate how organizations fulfill their data value vision and get the benefits of a data steward. In some organizations, data stewards are subject matter experts, owning and managing a discrete data subject area. In others, functional data stewards focus on a line of business and affiliate with the business objectives of the department. They may also oversee discrete business processes and oversee multiple data domains or application/systems.

More often, data stewards are IT centric, assigned to systems that generate the data they manage. Project-driven data stewards end their role when the project ends. In all these cases, ownership boundaries should be clear with defined business rules and usage environments.

In some scenarios, a project may devolve into poor data decisions, even if a data steward was involved initially. If a data steward goes off a project before the project is complete and without established data quality boundaries, it disrupts continuity. For example, project team members could unintentionally manipulate data while

generating a report to support business decisions, and in doing so create inaccuracies.

Data is often not reusable or scalable enterprise-wide, which does not help achieve broader business goals such as customer retention. With a solid foundation of the three components, you can reach the desired outcome under any data stewardship model.

## DISTRIBUTED DATA STEWARDSHIP MODEL IMPLEMENTATION

Similar to data stewardship models, distributed data stewardship models also warrant the use of the three core components. Key centralized IT functions such as data infrastructure, management and governance of data; appropriate data access for users; and regulatory compliance are potentially rearranged over time.

When implementing a distributed data stewardship model, these phases may be one of the following:

- Centralized IT and governance team with distributed data access management to a line of business for everyday user access.
- Centralized IT and decentralized governance team with distributed data access management, enabling domain-based access across the organization.
- Decentralized IT and governance team with distributed data access management.

**In this handbook:**

Data stewardship: Essential to data governance strategies

Enhance data governance with distributed data stewardship

Development of a data stewardship framework starts by selecting a model framework that best aligns with your organization. Then, define the core components of an operating model, and develop a transition plan to incrementally migrate to the target operating model. This will assist data governance, data sharing, and regular use of relevant data, while improving an organization's data ecosystems.