Introduction to Data Governance

Amanda Sackett

Abstract

In today's age, the generation of large amounts of both structured and unstructured data has expanded at an exponential rate, along with the complexity of data ecosystems. Key data sources are increasing in both size and volume, and the way that data is captured and assessed is shifting from strictly on-premise databases to cloud technology. Data has become critically valuable to every industry and department whether it be financial services, data science and analytics teams, sales and marketing, or healthcare. Organizations are becoming more reliant on data to run day-to-day operations and drive decision-making. In order to keep pace with the ever-expanding wave of source systems, digitization, and Big Data, organizations are making Data Governance an increasingly significant institution that is vital to both business and information technology (IT) strategies. Through the use of Data Governance frameworks, companies can effectively mitigate risk and maximize the effective use of data through clearly defined policies and procedures, streamlined processes, and active management of their data stores to ensure quality. This, in turn, allows for organizations to establish trustworthy data while delivering results with speed and efficiency. Many organizations are adopting Data Governance programs in order to manage their data ecosystems and while approaches may differ, successful programs share key similarities this paper seeks to discuss. Through the implementation of tailored Data Governance operating models, organizations can integrate policies and procedures into effective programs. Doing so will improve data accuracy, timeliness, and quality to allow organizations to transform data into information, and information into insight.

Keywords: Data Governance; Data Management; Data Governance Frameworks; Data Analytics; Data Quality; Data Stewardship; Data Remediation; Big Data; Data Lineage; Data Compliance

1. Introduction

Data has become the lifeblood in today's world – essential to running key company operations, driving decisions, and informing responses in almost all industries. As its scope expands, so too must the organizations that manage it. The ability to effectively create policies, procedures, and defined responsibilities surrounding this growing reality will lead to sustainable success in any organization. This success is realized by reducing the operational costs of managing data, mitigating the risks associated with incomplete and inaccurate data, and enabling the utilization of data as information to drive decisions.

Many organizations face common data challenges. These might include inconsistently defining data attributes across business units, struggling to find the data needed to quickly respond to regulatory requirements, not being able to create timely and consistent reporting for leadership, and an overall lack of understanding of enterprise data available to employees. By overcoming these challenges, organizations can achieve benefits such as increased organizational insight and visibility into the depth, profitability, and weaknesses of their customer relationships. To address these common challenges, organizations must adopt a Data Governance program, which is the formal orchestration of people, processes, and technology to enable an organization

to leverage its data as an enterprise asset. By proactively governing their data, organizations and industries are better able to achieve their strategies and objectives by decreasing the costs of managing data and increasing their ability to capitalize data as an asset. While there are variations of Data Governance frameworks, there are three universal pillars: policies and standards, processes, and roles and responsibilities.

When beginning to face the challenges presented by a lack of Data Governance, it is important to understand the fundamental difference between Data Governance and Data Management. These two terms are often interchanged when addressing data strategy but, while the two are interconnected, they are not the same. Data Governance is the high-level oversight of data providing the policies and procedures required to deliver complete, accurate, consistent, and timely data. Data Management[1], on the other hand, is the day-to-day execution of Data Governance policies, and includes the processes people follow as well as the tools and technologies they need to help them comply with the mandates of the Data Governance program. Data Management cannot succeed without Data Governance, which is why high-quality implementation of a Data Governance program is crucial to any long-term success. As the volume and complexity of data increases, so does the gap in Data Management capabilities when there is a continued absence of Data Governance. A Data Governance program should enforce the integration of strategy, standards, policies, and communication to deliver optimally. Regardless of the size of an organization's data scope, preparedness is crucial to ensure the quality delivery of data. This paper will explore the key cornerstones of any Data Governance program, the relevance of these programs, and the key advantages of implementing an effective Data Governance program into any organization.

"As the volume and complexity of data increases exponentially, so does the gap in Data Management capabilities with the continued absence of Data Governance."

2. Discussion

2.1. Motivations & Use Cases for Data Governance

Data is now a pervasive and mandatory aspect in any organization; thus, the importance of leadership in its understanding and investment cannot be downplayed. High-quality data is an asset which can be utilized to achieve a multitude of goals. Proper governance and management of data will ensure accuracy. This will not only manage risks but allow organizations to transform opportunities into advantages. By integrating Data Governance into an organization's corporate structure, organizations can increase trust in their data, improve efficiency to better utilize resources, and provide futureproofing to allow for the introduction of innovative technologies and data tools that will enhance data utilization.

An organization needs to know what it is investing in, how much the investment costs, and the value it brings, along with other metrics that drive decision-making. By having strong Data Governance policies and procedures, enterprise data tied to strategic initiatives can be effectively utilized to measure risk and returns while still controlling other costs. Highly accurate data will also lead to meaningful insight through clarity of reporting and analytics which will guide organizations to maximize profit. By eliminating errors from incorrect data, reducing

---

[1] *Data Management is the administrative process by which data is acquired, validated, stored, protected and processed to ensure accessibility, reliability, and timeliness of data.*

duplicative work for data feeds/stores, and reducing manual reconciliation or data manipulation, Data Governance will help to minimize costs by reducing the amount of effort needed in correcting and maintaining data stores. Overall, Data Governance will eliminate the need for having to divert internal resources to these tasks or having to hire outside of the organization to handle data cleanup.

One example of how Data Governance can help to work toward achieving these goals is by defining data policies that standardize key data attributes for greater data consistency and understanding across an organization. If an organization does not have standardized definitions for its data elements, data becomes convoluted and processes become inefficient. For something such as inventory management, having a standardized definition across the organization is crucial to ensuring product inventory is both complete and accurate, and that there are metrics available around sales, usage, age, etc. Ensuring this accuracy will allow for precise strategic planning and portfolio management downstream. Through data consistency, an organization can reduce cost and increase operational effectiveness by automating data reporting while informing strategic planning.

Another potential outcome of employing an effective Data Governance program is improved customer satisfaction resulting from data procedures and clean data. Today, people want clear visualizations of data that are digestible, and delivered quickly and regularly. In the healthcare industry, for example, this might come in the form of patient experience (PX) dashboards to display health facility and provider performance statistics along with various PX metrics that will inform on improvement of care.

A Data Governance program can also empower the creation of business dashboards in order to gauge opportunity streams or customer profitability profiles. This is done by leveraging consolidated data from across the organization which helps to recognize pain points as well as identify opportunity pipelines. To achieve this, an organization's data structures must be seamless across platforms to allow the data to be linked. Having clear policies and procedures in place will make an organization's data easy to manipulate and leverage into consolidated reporting and analytics platforms to provide up-to-date reporting. This will increase organizational capacity, alleviating the need for data cleanup and allowing an organization to effectively utilize data visualization and maximize understanding of its data. This provides insight on organizational direction, performance, and profits, which ultimately will inform downstream decision-making.

A Data Governance program can empower your organization to achieve a holistic view of your customers.

Data Governance policies and procedures ensure that the combination of data across channels and profiles can be leveraged in sales, customer service, and strategic planning. A Data Governance program can empower your organization to achieve a holistic view of your customers' relationship across the organization.

In addition to facilitating quality data visualization, which broadens the scope of data understanding, Data Governance also allows for the consolidation of a user's experiences. As data and systems expand and merge, bringing data together across systems becomes challenging. Data Governance programs support standardization across these systems which allows for a more seamless user experience. With technology platforms like web portals and applications becoming more relevant, a consolidated User Experience/Customer Experience (UX/CX) will improve overall experience and efficacy. Whether this is merging multiple login systems to work under one centralized hub or improving customer support and relationship managers (RMs), Data Governance can streamline messy systems and help with the marketing sales interactions. The evidence and motivation toward implementing a Data Governance program is overwhelming. Getting started in implementing

these programs can be a daunting task; however, the first step is choosing a framework that will guide how policies are created, procedures are implemented, and how the structure of the teams that will lead the program are determined.

2.2.   *The Data Governance Frameworks*

Once an organization decides to begin implementing a Data Governance program, it is important for it to define the scope of its data and to choose the right framework that meets its needs. This will determine how successful both the implementation and ongoing maintenance of the program will be. An organization must minimize data risk and create effective countermeasures against potential threats, while maximizing the benefits from data usage. Setting objectives and goals is the first measure to ensure a Data Governance program will meet the requirements for organizational success. Once policies are created and the procedures are in motion, there must be a team that will effectively manage and maintain those rules post-implementation. The structure of this team will vary based on an organization's unique structure and needs. There are a few approaches to resourcing a Data Governance program; however, there is no right answer that works best. In a case study examining two comparable telemarketing firms[2], each firm chose differing frameworks and configurations of Data Governance based on factors unique to the organization, including goals and structure. While different approaches were taken, both firms had successful Data Governance programs, demonstrating there is no "one size fits all" solution when it comes to Data Governance frameworks.

There are three common frameworks that are utilized when executing a Data Governance program (**Figure 1**). The first is the highly centralized approach. This involves fully allocated resources whose sole role responsibility is the Data Governance program. This approach is typically seen in large organizations or those with complex governance and data issues that necessitates a team fully dedicated to managing data policies. The next framework is the decentralized (or federated) approach, which involves partially allocated resources with no fully allocated resources existing solely for Data Governance. Instead, resources have Data Governance roles alongside other job responsibilities. This approach is typically seen in smaller organizations or those in the infancy of a Data Governance initiative. The last framework is the hybrid approach. This approach involves a mix of fully allocated and partially allocated resources in which some resources are fully dedicated to managing the program while others are partially allocated to the program alongside other responsibilities. The typical arrangement for this method is for the Data Governance Office to be instantiated with fully dedicated resources while a Steering Committee and Working Group are comprised of partially dedicated resources. As stated before, the approach selected by an organization will be dependent on an organization's unique attributes and needs.

The framework that has been selected will inform the roles that must be assigned to carry out the various responsibilities. Defining these roles and responsibilities is a critical aspect that will ensure the smooth running of any program. Core roles and groups include executive sponsors and stakeholders, a Steering Committee, an Office of Data Governance, a Working Group, data stewards, data creators, and data consumers.

---

[2] Otto, Boris. *Organizing Data Governance: Findings from the Telecommunications Industry and Consequences for Large Service Providers*. Communications of the Association for Information Systems. Volume 29, August 2011. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=3610&context=cais

In each of the frameworks, the data creators, stewards, and consumers will lay down the groundwork and maintain the underlying data for an organization. First, data creators will enter or extract data into/from a source system. The data stewards will then ensure the quality of that

Figure 1. Data Governance Frameworks
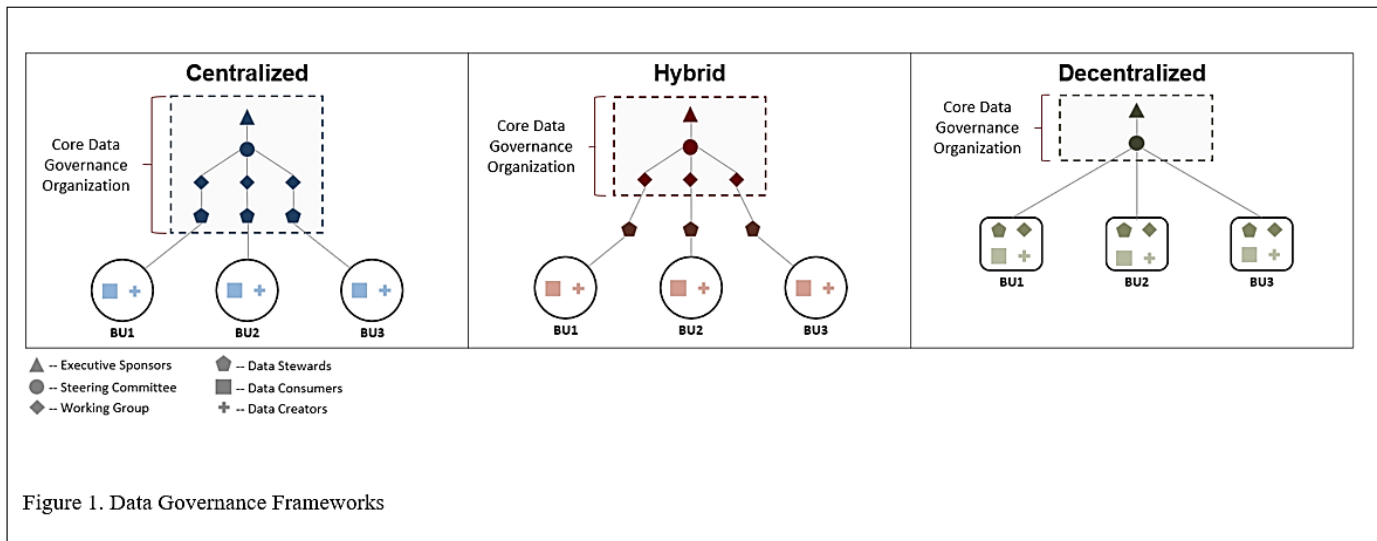
data by performing data review and cleansing activities. These individuals will serve as liaisons to business units regarding governance policy and procedure items, consolidation of systems, data consumers and originators, and IT support. The final aspect behind the data in an organization will be the data consumers who utilize the data for both operational systems, tasks, and analytics. As the data flows downstream, systems and employees will use this data for process support and organizational engagement.

While the data teams are critical, executive sponsorship and stakeholder buy-in are crucial for the success of any Data Governance program. These are the individuals who understand corporate strategies and working with them will ensure the program is built to align with these strategies. This will guarantee that stakeholders remain committed to the program and understand its overall business importance. The most common reason for failure in Data Governance initiatives is a lack of executive sponsorship. These sponsors are typically organization leads who will ensure funding and resources are available to execute the Data Governance program.

These players roll up into three governing bodies that oversee an organization's Data Governance program. The first is the Steering Committee, which is the core leadership for the Data Governance initiative. This committee will approve funding, budget, and resource allocation for strategic data projects. It also will ensure that there is alignment with corporate objectives and define the overall road map and priorities. The Steering Committee must be ready to assist in enforcement to guarantee that all individuals responsible for Data Governance activities are held accountable, while also looking inward to ensure its own adherence.

The next governing body is the Data Governance Working Group, which oversees the execution of the Data Governance program and much of the day-to-day data quality management. It provides the guidance and enforcement to drive change in the use and maintenance of data. These individuals will mainly focus on Data Governance execution through training, education, and promoting awareness to the organization's employees. This group will also serve as key stakeholders in the definition and implementation of policies and procedures.

Finally, there is the Office of Data Governance, which is comprised of chair members from both the Steering Committee and Data Governance Working Group. This office handles the overall management of Data Governance and will define the standards, metrics, and procedures for data quality across the organization and advise business and technical resources on data standards to ensure data quality. This group serves as the front line to all impacted parties affected by the deployment of a new data policy.

2.3.   Implementing Data Governance

After a framework is chosen and roles and responsibilities have been defined, the implementation of a Data Governance program can begin. The first step in developing a Data Governance program is to establish Data Governance policies that will govern and act as the backbone of the program. The policies will function as an internal regulation mechanism, one with which organizations must comply to effectively ensure they have the right resources, structure, and enforcement within their programs. When setting a Data Governance policy, organizations must first look inward to determine the current state of their corporate strategies. For example, they might ask what role their information plays in decision-making as they assess their performance against corporate strategies. With that in mind, existing data structures must be analyzed to determine where the pitfalls are in the current system that would hinder organizations from achieving their strategic goals, and what would need to change to enable their success. This frame of mind will help mold the creation of a policy specific to an organization's needs. By creating policies, the organization is creating a rulebook and statements defining how data will be managed and what its Data Governance program will entail.

Policies will be created at both the organizational and data element levels. The organizational-level policies, sometimes referred to as "anchor policies," address the overarching rules and structures with which organizations must comply. These rules lay the groundwork for how an organization will effectively manage its Data Governance, and how they apply to the organization as a whole. They also will set the stage for the overall program within an organization. The organizational-level policies might be setting a rule that simply states an organization will *have* a formalized Data Governance program and define the organizational framework, membership, authority, charter, and other such infrastructure elements to which it will adhere. Additional organizational-level policies might lay out the foundation for data practices including mandates for the creation and maintenance of data dictionaries for all company data including related metadata, lineage, and data quality rules. This kind of policy will define part of the work that the Data Governance program needs to support.

Lower-level data policies will cover the governing of the actual data and data elements. While anchor policies will set the rules of engagement surrounding Data Governance, the data element policies will work downstream to improve the quality, accuracy, completeness, and consistency of data. This is where the real change occurs. An example of this type of policy might include rules governing the appropriate use of a key field such as a status field. A policy governing status may provide guidance on appropriate values (i.e., quality, accuracy), whether to make the field required (i.e., completeness), and that these values be leveraged across all systems that use a status field (i.e., consistency). This policy, when set, will act as the rulebook by which subsequent procedures and processes are executed. By laying down this foundation, it will establish the basis on which the remainder of the program and procedures rely. Without a policy in place, there is no structure or driving force behind any future actions, which can lead to the failure of a Data Governance program.

"While there are variations of Data Governance frameworks, there are three universal pillars: policies and standards, processes, and roles and responsibilities."

As policies for a Data Governance program are finalized, an organization must then lay out the procedures to implement that policy or correct existing procedures that are counter to the newly created policy. For example, a Data Governance policy might require all new customer records have a unique identifier (e.g., Tax ID) populated with nine numeric digits. From there, procedures to certify this policy is followed could include data entry, exception reviews, duplication management, and any others needed across all systems in which the concept of customer originates and/or is used. The Data Governance organization would lead the creation of technical tools (e.g., front-end entry screens, required fields, drop-down lists) as well as resource responsibilities (e.g., weekly data reviews, training, communication) to efficiently execute said procedures. This execution and enforcement of policies is critical for ensuring trustworthy data. While technology is a valuable tool, policies resulting in training, Change Management, and communication are also viable (and often cost-effective) tactics that should be implemented.

Another important practice that should underlie any program is data quality reporting to prove that the Data Governance program is successful and useful. By periodically measuring data quality from the start of the Data Governance Program, organizations can provide insights

Key metrics and key performance indicators (KPIs) for Data Governance programs[3] are:

Program Performance

- Number of supported groups and users
- Issue Metrics: Number of issues, resolution time, issue status
- Qualitative engagement scores
- Data Steward KPI: Number of items for review, resolution time

Data Quality

- Record error / review required alerts
- Duplicate record check, orphan record alerts
- Data Integrity Metrics: Record completeness, field uniqueness, data consistency, and data connection density

Business Value

- Regulatory compliance status
- Risk mitigation (i.e., costs avoided)
- Cost Savings: Resource time, software costs, maintenance costs
- Value Creation: Lift in cross-sell / up-sell revenue, customer retention, customer satisfaction, resource allocation

Figure 2. Key metrics and key performance indicators will ensure a Data Governance program maximized program performance, data quality, and business value

into how the program is improving data quality, as well as highlight areas that need more improvement. Procedures must also be put in place to account for issue management. Constant monitoring and management of the program while intermittently making tactical fixes will ensure the data retains the appropriate level of quality and consistency. These fixes might include Data Governance procedures such as creating data definitions, defining attributes, amending attributes, and retiring attributes. Issue management will also entail revising the procedures so that a Data Governance program continues to evolve to best meet the organization's needs. Overall, a successful program will have policies and procedures working in tandem, with an organized team that will implement and enforce the Data Governance program.

[3] Karel, Rob. *Monitoring Data with Data Monitoring Tools*. Informatica Blog. 2 January 2014. https://blogs.informatica.com/2014/01/02/data-governance-measure-monitor-processes-open-a-window-to-data-and-its-value/

Once an organization has chosen the Data Governance framework that will best suit its needs and initiated Data Governance deployment, it should then define the methods for ongoing monitoring and maintenance of the program. As discussed above, adding accountability to resource roles and responsibilities is key to ensuring the ownership and implementation of Data Governance policies. By doing this, it will guarantee procedures become a part of day-to-day operations. Furthermore, actively monitoring the performance of the Data Governance program both in terms of data quality improvements and the attainment of business values (e.g., lower costs, operational efficiency, increased sales) is critical to continually improving an organization's Data Governance operating model. Creating reporting assets, either through the purchase of a Data Governance tool or through an internal build, is one way to address these monitoring tasks. Analytics dashboards, scorecards, and status updates provide managers and operators with insights into the overarching performance of a Data Governance platform.

At a Financial Services client, Kenway implemented a Client Data Governance Program by:

1. Developing policies to govern how client data was captured and which identifier values were required.

2. Designing the data remediation steps needed to cleanse historical data and processes to ensure ongoing data integrity.

3. Building the business case for an enterprise data management program for client data.

4. Deploying a Client Data Hub allowing for improved risk rating, lower manual work, and a clear view into a client's holdings across the organization.

It is of note that a large component of implementation is Change Management. There is a need to align both people and processes to strategic initiatives. By applying tenets of Change Management[4], an organization can ensure there is understanding and buy-in from both stakeholders and all impacted resources which will aid in the successful rollout of the program. By setting incremental goals and including team members who will be impacted by the new program, the organization will build momentum within the program, allow for feedback and increased participant buy-in, and create a sense of accountability in the successful rollout. As the process continues, there is a need for continuous evaluation to understand which components are successful and which are not. Measures to gain stakeholder trust include setting success measures followed by evaluation, recommendations for improvement, and broadcasting success. This trust and collaboration are critical as stakeholders will be the ones to advocate for and expand Data Governance programs throughout an organization. By utilizing Change Management and broadcasting success, an organization can minimize risks associated with significant culture shifts and guarantee governance thrives instead of falling to the wayside.

2.4.   Data Governance Maturity Curve

Overall, it is good practice to periodically assess the maturity of Data Governance in an organization. Quantifying and defining scope and ability lets an organization assess the state of its program, understand its data capabilities, identify vulnerabilities, and know which areas need improvement. The level of maturity and complexity of an organization's Data Governance monitoring capability, as well as its overall Data Governance program, helps to identify where it falls on the Data Governance Maturity Curve (**Figure 3**). When an organization's Data
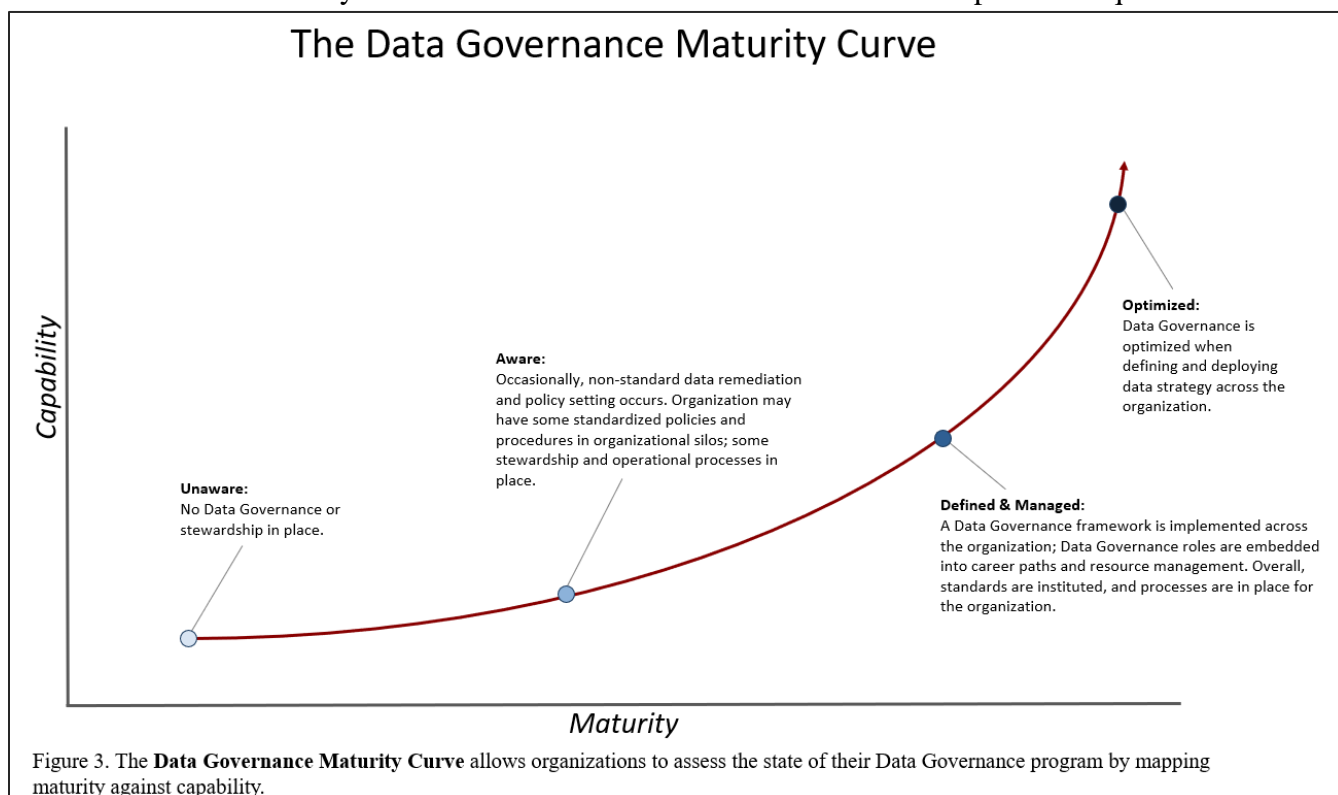
---

[4] Best practices of Change Management include acknowledging underlying culture, leadership commitment, building change capability, constant communication, and transparency.

Governance maturity is at the Unaware stage, there are no policies or procedures in place to ensure the enterprise-wide or even cross-functional quality of data. This often arises as organizations grow, while early-stage organizations can afford to allow distinct groups and business units to function separately. As the utilization of data across an organization increases, the organization moves to a higher level of maturity, the Aware stage, where there is some siloed Data Governance to support its growing capabilities. These organizations typically require independent efforts to create consistent data. For example, a sales team may require its users to enter an email address for all prospects in order to tie customer interactions to orders. This would likely require some manual reviews of records and addressing missing data on an ad-hoc basis without the appropriate Data Governance in place.

As an organization begins to formalize Data Governance, it moves into the Defined & Managed stage of the Maturity Curve. Here, standards and policies (similar to the email address example above) are in place across the organization which allows for a more robust connection of information across all systems. The formalization of these standards and policies requires the

The Data Governance Maturity Curve

Aware:
Occasionally, non-standard data remediation and policy setting occurs. Organization may have some standardized policies and procedures in organizational silos; some stewardship and operational processes in place.

Unaware:
No Data Governance or stewardship in place.

Optimized:
Data Governance is optimized when defining and deploying data strategy across the organization.

Defined & Managed:
A Data Governance framework is implemented across the organization; Data Governance roles are embedded into career paths and resource management. Overall, standards are instituted, and processes are in place for the organization.

Capability (y-axis) — *Maturity* (x-axis)

Figure 3. The **Data Governance Maturity Curve** allows organizations to assess the state of their Data Governance program by mapping maturity against capability.

creation of dedicated roles and responsibilities for Data Governance. As discussed, Data Governance frameworks can vary by organization, but they each fundamentally involve a core governance team and ties to resources' job descriptions. Finally, the organizations with the most mature Data Governance optimize their existing operating framework to allow for the enhancements of strategic assets throughout the organization. These companies factor in Data Governance as they deploy new services, product features, and incorporate new data sets. This allows them to quickly and efficiently integrate data into their current operations, minimizing the time to value and costs to manage the data.

Wherever on the maturity spectrum an organization may be, understanding the changes required to reach the next step will be vital to successful implementation. As the complexity of systems increase and an organization progresses along the maturity curve, it will find that it needs to address the shifting landscape of its organization which often comes in the form of

organizational change management. Whatever the method, as the organization evolves, it must continue to optimize its processes to realize the highest level of benefits and remain competitive on the market. By continuously assessing capability and utilizing a maturity curve, an organization can achieve a successful and long-lasting Data Governance program.

3. Conclusion

As the world continues to have access to more data, we will be faced with the challenge of data integration and expanding data structures. Data Governance is the key in exercising authority and control in the management of that data. As data has become increasingly relevant and imperative in decisions and policymaking, it is crucial for industries to integrate Data Governance as a key structure in their organization. Implementation of a Data Governance program will ensure high-quality information and help organizations with strategic decision-making to maximize profitability. Data Governance must be continually cultivated, and organizations must look to be innovative in adapting their Data Governance frameworks to best grow alongside evolving industries and technologies. Future studies can look to either build upon existing frameworks or create completely new methodologies in order to effectively manage and adapt to growing information and stay ahead of the curve. The data landscape is rapidly changing, and it is necessary for Data Governance to evolve along with it so that organizations continue to generate value and maintain sustainable competitive advantages. With technologies like artificial intelligence, machine learning, text mining, and predictive analytics on the forefront of business strategy, there is a multitude of opportunities to cultivate a Data Governance program and fully extract the maximum value from data enterprises. Coupled with focusing Data Governance policies to best achieve strategic goals, organizations that choose to adopt Data Governance as a standard will thrive.

Kenway Consulting is a management and technology consulting firm whose entire reason for existence is to help companies and their employees. Kenway strives to provide all clients with unmatched quality and service, and specializes in the areas of Technology Solution Delivery, Enterprise Program Leadership, and Information Insight.

This white paper was published as a reference for individuals and organizations embarking on or working to improve their Data Governance programs.

For additional information, please contact Kenway via email (info@kenwayconsulting.com) or visit our Data Governance Service page (linked here).