DAMA SUMMARY

# Data Management

***Data Management is the development, execution, and supervision of plans, policies, programs, and practices that deliver, control, protect, and enhance the value of data and information assets throughout their lifecycles.***

* *The primary driver for data management is to enable organizations to get value from their data assets, just as effective management of financial and physical assets enables organizations to get value from those assets.*
* *Information and knowledge hold the key to competitive advantage. Organizations that have reliable, high quality data about their customers, products, services, and operations can make better decisions than those without data or with unreliable data.*

## Data Management Goals

* *Understanding and supporting the information needs of the enterprise and its stakeholders, including customers, employees, and business partners*
* *Capturing, storing, protecting, and ensuring the integrity of data assets*
* *Ensuring the quality of data and information*
* *Ensuring the privacy and confidentiality of stakeholder data*
* *Preventing unauthorized or inappropriate access, manipulation, or use of data and information*
* *Ensuring data can be used effectively to add value to the enterprise*

## What is data?

* *Data is a means of representation. It stands for things other than itself (Chisholm, 2010).*
* *Data is both an interpretation of the objects it represents and an object that must be interpreted (Sebastian-Coleman, 2013).*
* *This is another way of saying that we need context for data to be meaningful. Context can be thought of as data’s representational system; such a system includes a common vocabulary and a set of relationships between components. If we know the conventions of such a system, then we can interpret the data within it. These conventions are often documented in a specific kind of data referred to as Metadata.*
* *Even within a single organization, there are often multiple ways of representing the same idea. Hence the need for Data Architecture, modeling, governance, and stewardship, and Metadata and Data Quality management, all of which help people understand and use data. Across organizations, the problem of multiplicity multiplies. Hence the need for industry-level data standards that can bring more consistency to data.*

Much ink has been spilled over the relationship between data and information. Data has been called the “raw material of information” and information has been called “data in context”. Often a layered Pyramid is used to describe the relationship between data (at the base), information, knowledge, and wisdom (at the very top).

An asset is an economic resource, that can be owned or controlled, and that holds or produces value. Assets can be converted to money. Data is widely recognized as an enterprise asset.

Many organizations identify themselves as ‘data-driven’. Businesses aiming to stay competitive must stop making decisions based on gut feelings or instincts, and instead use event triggers and apply analytics to gain actionable insight.

## Data Management Principles

* *Data is an asset with unique properties: Data is an asset, but it differs from other assets in important ways that influence how it is managed. The most obvious of these properties is that data is not consumed when it is used, as are financial and physical assets.*
* *The value of data can and should be expressed in economic terms;* *A primary challenge to data asset valuation is that the value of data is contextual (what is of value to one organization may not be of value to another) and often temporal (what was valuable yesterday may not be valuable*

*today).*

* *Managing data means managing the quality of data*
* *It takes Metadata to manage data: Managing any asset requires having data about that asset (number of employees, accounting codes, etc.). The data used to manage and use data is called Metadata. Because data cannot be held or touched, to understand what it is and how to use it requires definition and knowledge in the form of Metadata. Metadata originates from a range of processes related to data creation, processing, and use, including architecture, modeling, stewardship, governance, Data Quality management, systems development, IT and business operations, and analytics.*
* *Data management is cross-functional; it requires a range of skills and expertise*
* *Data management requires an enterprise perspective*
* *Data management is lifecycle management: Data has a lifecycle and managing data requires managing its lifecycle. Because data begets more data, the data lifecycle itself can be very complex. Data management practices need to account for the data lifecycle.*
* *Data management requirements must drive Information Technology decisions*

Data Quality

As importantly, poor quality data is simply costly to any organization. Estimates differ, but experts think

organizations spend between 10-30% of revenue handling data quality issues. IBM estimated the cost of poor quality data in the US in 2016 was $3.1 Trillion.

Many of the costs of poor quality data are hidden, indirect, and therefore hard to measure. Others, like fines, are direct and easy to calculate. Ensuring that data is of high quality is central to data management. Organizations manage their data because they want to use it. If they cannot rely on it to meet business needs, then the effort to collect, store, secure, and enable access to it is wasted. To ensure data meets business needs, they must work with data consumers to define these needs, including characteristics that make data of high quality.

Planning for better data: As stated in the chapter introduction, deriving value from data does not happen by accident. It requires planning in many forms. It starts with the recognition that organizations can control how they obtain and create data. If they view data as a product that they create, they will make better decisions about it throughout its lifecycle.

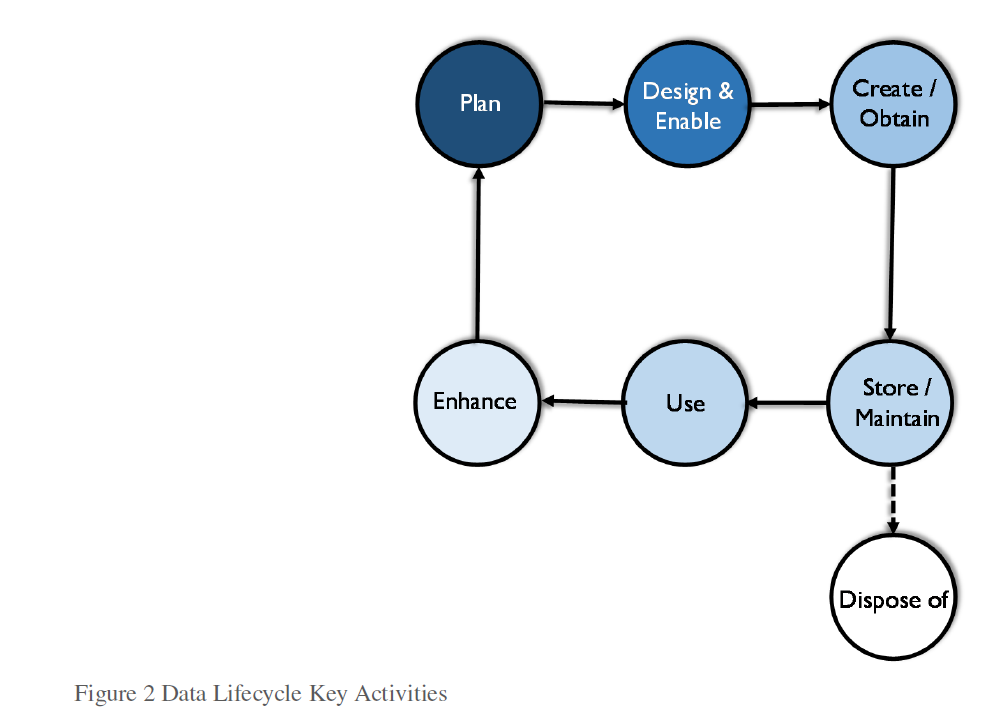
Meta Data

* Organizations require reliable Metadata to manage data as an asset. Metadata in this sense should be understood comprehensively. It includes not only the business, technical, and operational Metadata but also the Metadata embedded in Data Architecture, data models, data security requirements, data integration standards, and data operational processes.
* Data is abstract. Definitions and other descriptions of context enable it to be understood. They make data, the data lifecycle, and the complex systems that contain data comprehensible.

Enterprise perspective

* Managing data requires understanding the scope and range of data within an organization. Data is one of the ‘horizontals’ of an organization. It moves across verticals, such as sales, marketing, and operation

Data Lifecycle



Data Quality\MetaData Quality and security must be managed throughout the data lifecycle.

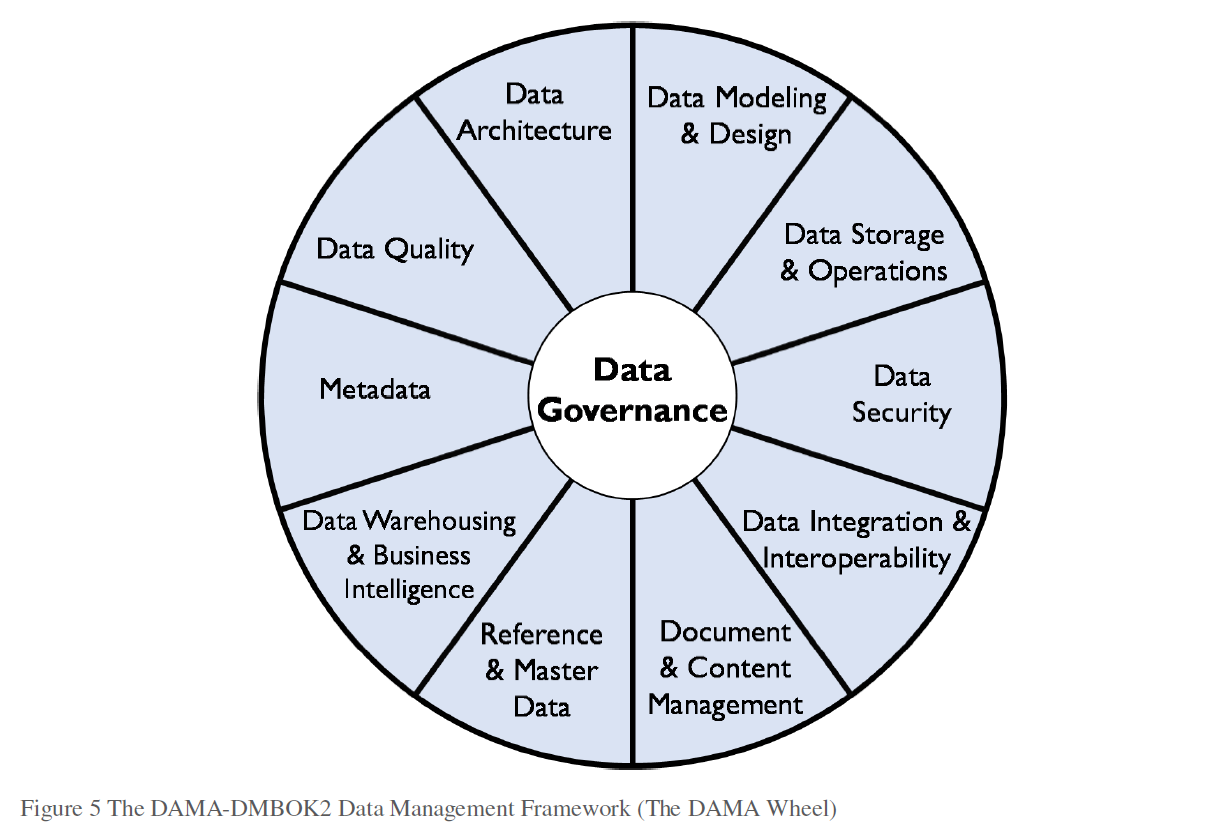
Data and Risk

Data not only represents value, it also represents risk. Low quality data (inaccurate, incomplete, or out-of-date) obviously represents risk because its information is not right. But data is also risky because it can be misunderstood and misused.

Data Strategy

A data strategy should include business plans to use information to competitive advantage and support enterprise goals. Data strategy must come from an understanding of the data needs inherent in the business strategy: what data the organization needs, how it will get the data, how it will manage it and ensure its reliability over time, and how it will utilize it.

* Data Management program objectives that are SMART (specific, measurable, actionable, realistic, time-bound)



Data Governance

Data Governance (DG) is defined as the exercise of authority and control (planning, monitoring, and

enforcement) over the management of data assets. All organizations make decisions about data,

regardless of whether they have a formal Data Governance function. Those that establish a formal Data

Governance program exercise authority and control with greater intentionality (Seiner, 2014). Such organizations are better able to increase the value they get from their data assets.

Goals:

1. Enable an organization to manage its data as an asset.

2. Define, approve, communicate, and implement principles, policies, procedures, metrics, tools, and

responsibilities for data management.

3. Monitor and guide policy compliance, data usage, and management activities.

The most common driver for data governance is often regulatory compliance, especially for heavily regulated industries, such as financial services and healthcare. Responding to evolving legislation requires strict data governance processes. The explosion in advanced analytics and Data Science has created an additional driving force.

Data governance is not an end in itself. It needs to align directly with organizational strategy. The more clearly it

helps solve organizational problems, the more likely people will change behaviors and adopt governance practices.

Reducing Risk

o **General risk management**: Oversight of the risks data poses to finances or reputation, including

response to legal (E-Discovery) and regulatory issues.

o **Data security**: Protection of data assets through controls for the availability, usability, integrity,

consistency, auditability and security of data.

o **Privacy**: Control of private / confidential / Personal Identifying Information (PII) through policy

and compliance monitoring.

• Improving Processes

o **Regulatory compliance**: The ability to respond efficiently and consistently to regulatory

requirements.

o **Data quality improvement**: The ability to contribute to improved business performance by

making data more reliable.

o **Metadata Management**: Establishment of a business glossary to define and locate data in the

organization; ensuring the wide range of other Metadata is managed and made available to the

organization.

o **Efficiency in development projects**: SDLC improvements to address issues and opportunities in

data management across the organization, including management of data-specific technical debt

through governance of the data lifecycle.

o **Vendor management**: Control of contracts dealing with data, such as cloud storage, external

data purchase, sales of data as a product, and outsourcing data operations.

* **Strategy**: Defining, communicating, and driving execution of Data Strategy and Data Governance Strategy
* **Policy**: Setting and enforcing policies related to data and Metadata management, access, usage, security, and quality
* **Standards and quality**: Setting and enforcing Data Quality and Data Architecture standards
* **Oversight**: Providing hands-on observation, audit, and correction in key areas of quality, policy, and data management (often referred to as *stewardship)*
* **Compliance**: Ensuring the organization can meet data-related regulatory compliance requirements
* **Issue management**: Identifying, defining, escalating, and resolving issues related to data security, data access, data quality, regulatory compliance, data ownership, policy, standards, terminology, or data governance procedures
* **Data management projects**: Sponsoring efforts to improve data management practices
* **Data asset valuation**: Setting standards and processes to consistently define the business value of data assets

To achieve this overall goal, a DG program must be:

* **Sustainable**: The DG program needs to be ‘sticky’. DG is not a project with a defined end; it is an ongoing process that requires organizational commitment. DG necessitates changes in how data is managed and used. This does not always mean massive new organizations and upheaval. It does mean managing change in a way that is sustainable beyond the initial implementation of any data governance component. Sustainable data governance depends on business leadership, sponsorship, and ownership.
* **Embedded**: DG is not an add-on process. DG activities need to be incorporated into development methods for software, use of data for analytics, management of Master Data, and risk management.
* **Measured**: DG done well has positive financial impact, but demonstrating this impact requires understanding the starting point and planning for measurable improvement.

*Data-centric Organization*

A data-centric organization values data as an asset and manages data through all phases of its lifecycle, including

project development and ongoing operations. To become data-centric, an organization must change the way it

translates strategy into action. Data is no longer treated as a by-product of process and applications. Ensuring data

is of high quality is a goal of business processes. As organizations strive to make decisions based on insights gained

from analytics, effective data management becomes a very high priority.

People tend to conflate data and information technology. To become data-centric, organizations need to think

differently and recognize that managing data is different from managing IT. This shift is not easy. Existing culture,

with its internal politics, ambiguity about ownership, budgetary competition, and legacy systems, can be a huge

obstacle to establishing an enterprise vision of data governance and data management.

*Data Governance Organization*

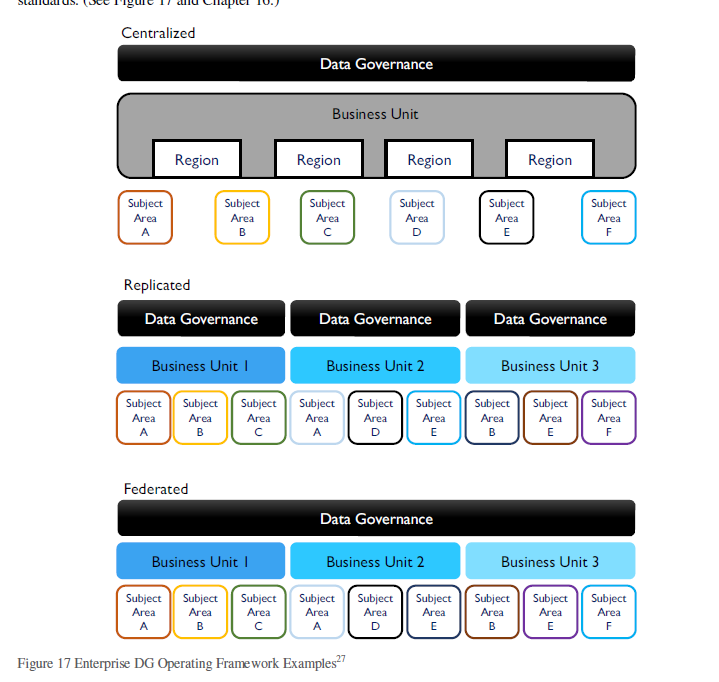
The core word in governance is *govern*. Data governance can be understood in terms of political governance. It

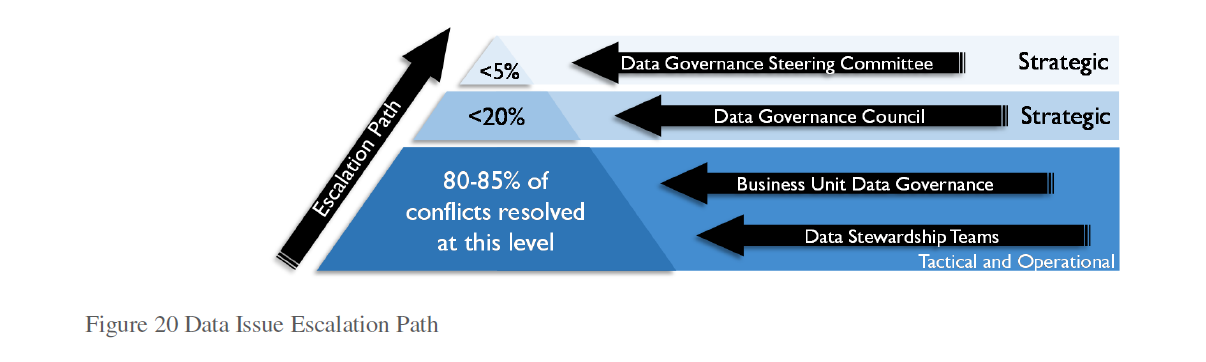
includes

* legislative-like functions (defining policies, standards, and the Enterprise Data Architecture),
* judicial-like functions (issue management and escalation), and
* executive functions (protecting and serving, administrative responsibilities).

*Data Governance Operating Model Types*

In a centralized model, one Data Governance organization oversees all activities in all subject areas. In a replicated model, the same DG operating model and standards are adopted by each business unit. In a federated model, one Data Governance organization coordinates with multiple Business Units to maintain consistent definitions and standards.





*Data Stewardship* is the most common label to describe accountability and responsibility for data and processes

that ensure effective control and use of data assets.However, in most cases, data stewardship activities will focus on some, if not all, of the following:

* Creating and managing core Metadata: Definition and management of business terminology, valid data values, and other critical Metadata. Stewards are often responsible for an organization’s Business Glossary, which becomes the system of record for business terms related to data.
* Documenting rules and standards: Definition/documentation of business rules, data standards, and data quality rules. Expectations used to define high quality data are often formulated in terms of rules rooted in the business processes that create or consume data. Stewards help surface these rules in order to ensure that there is consensus about them within the organization and that they are used consistently.
* Managing data quality issues: Stewards are often involved with the identification and resolution of data related issues or in facilitating the process of resolution.
* Executing operational data governance activities: Stewards are responsible for ensuring that, day-to day and project-by-project, data governance policies and initiatives are adhered to. They should influence decisions to ensure that data is managed in ways that support the overall goals of the organization.

Depending on the complexity of the organization and the goals of its DG program, formally appointed Data

Stewards may be differentiated by their place within an organization, by the focus of their work, or by both. For

example:

* **Chief Data Stewards** may chair data governance bodies in lieu of the CDO or may act as a CDO in a virtual (committee-based) or distributed data governance organization. They may also be Executive Sponsors.
* **Executive Data Stewards** are senior managers who serve on a Data Governance Council.
* **Enterprise Data Stewards** have oversight of a data domain across business functions.
* **Business Data Stewards** are business professionals, most often recognized subject matter experts, accountable for a subset of data. They work with stakeholders to define and control data.
* **A Data Owner** is a business Data Steward, who has approval authority for decisions about data within their domain.
* **Technical Data Stewards** are IT professionals operating within one of the Knowledge Areas, such as Data Integration Specialists, Database Administrators, Business Intelligence Specialists, Data Quality Analysts or Metadata Administrators.
* **Coordinating Data Stewards** lead and represent teams of business and technical Data Stewards in discussions across teams and with executive Data Stewards. Coordinating Data Stewards are particularly important in large organizations.