IF3140 Database System

Data Governance

SEMESTER I 2024/2025

KNOWLEDGE & SOFTWARE ENGINEERING

Source

DAMA DMBOK: Data Management Body of Knowledge, 2nd Edition, Technics Publications, 2017

- Chapter 3 Data Governance

Outline

- Introduction
- Goals and Principles
- Business Drivers
- Data-centric Organization
- Data Governance Organization
- Data Governance Operating Model Types
- Data Stewardship

- Other issues: Data Policies & Data Asset Valuation
- Data Governance Activities
- Data Governance Metrics
- Tools and Techniques
  - Data Governance Scorecards
  - Data Governance Maturity Model

Introduction

Data Governance is defined as *the exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets*
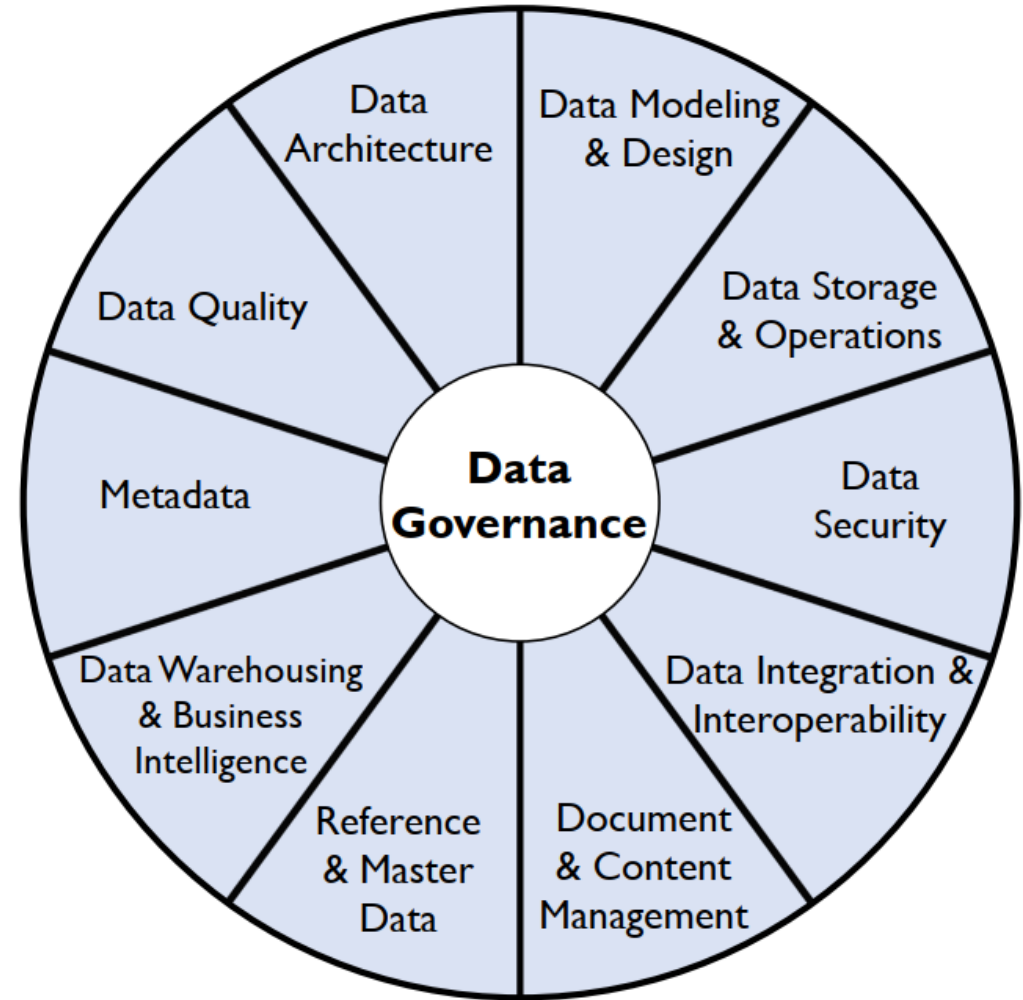
DAMA DMBOK2 Data Management Framework
Copyright @2017 by DAMA International

Introduction

- The **purpose** of Data Governance is to ensure that data is managed properly, according to policies and best practices (Ladley, 2012)

- Data Governance focuses on how decisions are made about data and how people and processes are expected to behave in relation to data

D.A.M.A DMBOK Data Governance Context Diagram

Data Governance and Stewardship

Definition: The exercise of authority, control, and shared decision-making (planning, monitoring, and enforcement) over the management of 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.
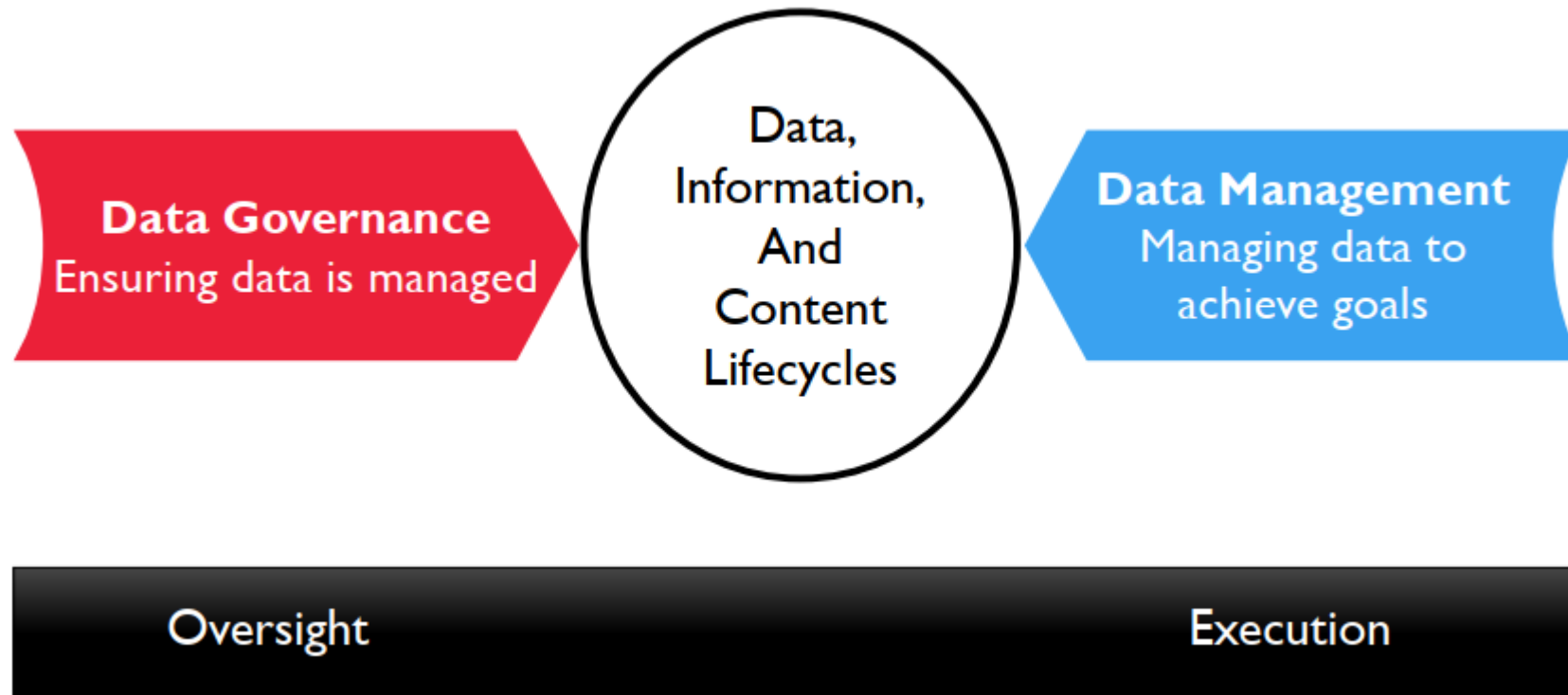3. Monitor and guide policy compliance, data usage, and management activities.

Business Drivers

Inputs:
- Business Strategies & Goals
- IT Strategies & Goals
- Data Management and Data Strategies
- Organization Policies & Standards
- Business Culture Assessment
- Data Maturity Assessment
- IT Practices
- Regulatory Requirements

Activities:
1. **Define Data Governance for the Organization (P)**
   1. Develop Data Governance Strategy
   2. Perform Readiness Assessment
   3. Perform Discovery and Business Alignment
   4. Develop Organizational Touchpoints
2. **Define the Data Governance Strategy (P)**
   1. Define the Data Governance Operating Framework
   2. Develop Goals, Principles, and Policies
   3. Underwrite Data Management Projects
   4. Engage Change Management
   5. Engage in Issue Management
   6. Assess Regulatory Compliance Requirements
3. **Implement Data Governance (O)**
   1. Sponsor Data Standards and Procedures
   2. Develop a Business Glossary
   3. Co-ordinate with Architecture Groups
   4. Sponsor Data Asset Valuation
4. **Embed Data Governance (C,O)**

Deliverables:
- Data Governance Strategy
- Data Strategy
- Business / Data Governance Strategy Roadmap
- Data Principles, Data Governance Policies, Processes
- Operating Framework
- Roadmap and Implementation Strategy
- Operations Plan
- Business Glossary
- Data Governance Scorecard
- Data Governance Website
- Communications Plan
- Recognized Data Value
- Maturing Data Management Practices

Suppliers:
- Business Executives
- Data Stewards
- Data Owners
- Subject Matter Experts
- Maturity Assessors
- Regulators
- Enterprise Architects

Participants:
- Steering Committees
- CIO
- CDO / Chief Data Stewards
- Executive Data Stewards
- Coordinating Data Stewards
- Business Data Stewards
- Data Governance Bodies
- Compliance Team
- DM Executives
- Change Managers
- Enterprise Data Architects
- Project Management Office
- Governance Bodies
- Audit
- Data Professionals

Consumers:
- Data Governance Bodies
- Project Managers
- Compliance Team
- DM Communities of Interest
- DM Team
- Business Management
- Architecture Groups
- Partner Organizations

Technical Drivers

Techniques:
- Concise Messaging
- Contact List
- Logo

Tools:
- Websites
- Business Glossary Tools
- Workflow Tools
- Document Management Tools
- Data Governance Scorecards

Metrics:
- Compliance to regulatory and internal data policies.
- Value
- Effectiveness
- Sustainability

(P) Planning, (C) Control, (D) Development, (O) Operations

INSTITUT TEKNOLOGI BANDUNG · 1920

KSE
KNOWLEDGE & SOFTWARE ENGINEERING

Data Governance and Data Management

Data governance represents an *inherent separation of duty between oversight and execution.*
Data governance ensures data is properly managed without directly executing data management.

Business Drivers

- 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

Business Drivers

- Reducing risks
  - **General risk management**: Oversight of the risks data poses to finances or reputation, incl. legal and regulatory issues
  - **Data security**: Protection of data assets through controls for the availability, usability, integrity, consistency, auditability and security of data
  - **Privacy**: Control of private/confidential/Personal Identifying Information (PII) through policy and compliance monitoring

KNOWLEDGE & SOFTWARE ENGINEERING

Business Drivers

- Improving processes
    - **Regulatory compliance**: The ability to respond efficiently and consistently to regulatory requirements.
    - **Data quality improvement**: The ability to contribute to improved business performance by making data more reliable.
    - **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.
    - **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.
    - **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.

Data Governance Goals

- To enable an organization to manage data as an asset.

- Provides the principles, policy, processes, framework, metrics, and oversight to manage data as an asset and to guide data management activities at all levels

- Monitor and guide policy compliance, data usage, and management activities

Data Governance Goals

- To achieve the goals, a data governance program must be:
  - **Sustainable**: Data governance is an ongoing process that requires organizational commitment. Sustainable data governance depends on business leadership, sponsorship, and ownership.
  - **Embedded**: Data governance activities need to be incorporated into development methods for software, use of data for analytics, management of Master Data, and risk management.
  - **Measured**: Data governance done well has positive financial impact, but demonstrating this impact requires understanding the starting point and planning for measurable improvement.

Data Governance Principles

- **Leadership and strategy**: Successful Data Governance starts with visionary and committed leadership. Data management activities are guided by a data strategy that is itself driven by the enterprise business strategy.

- **Business-driven**: Data Governance is a business program, and, as such, must govern IT decisions related to data as much as it governs business interaction with data.

- **Shared responsibility**: Across all Data Management Knowledge Areas, data governance is a shared responsibility between business data stewards and technical data management professionals.

- **Multi-layered**: Data governance occurs at both the enterprise and local levels and often at levels in between.

- **Framework-based:** Because data governance activities require coordination across functional areas, the Data Governance program must establish an operating framework that defines accountabilities and interactions.

- **Principle-based**: Guiding principles are the foundation of Data Governance activities, and especially of Data Governance policy.

Data-centric organization

- A data-centric organization values data as an asset and manages data through all phases of its lifecycle
    - Ensuring data is of high quality is a goal of business processes
    - Organizations need to recognize that managing data is different from managing IT
- Some shared principles of data-centric organization
    - Data should be managed as a corporate asset
    - Data management best practices should be incented across the organization
    - Enterprise data strategy must be directly aligned with overall business strategy
    - Data management processes should be continuously improved

KNOWLEDGE & SOFTWARE ENGINEERING

Data Governance Organization

Includes:

- **legislative**-like functions (defining policies, standards, and the Enterprise Data Architecture)

- **judicial**-like functions (issue management and escalation)

- **executive** functions (protecting and serving, administrative responsibilities)
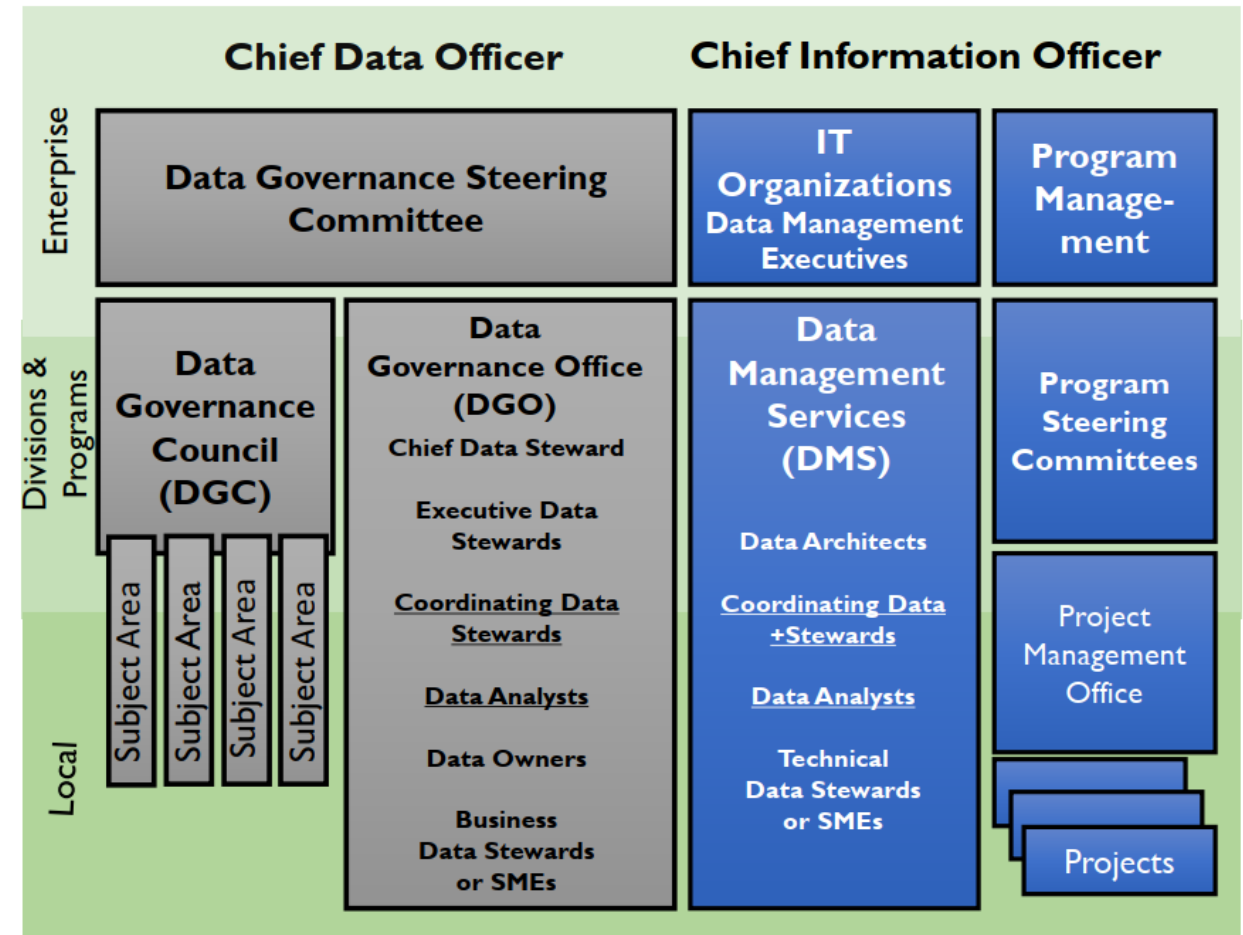
Data Governance Organization

Data Governance Body	Description
Data Governance Steering Committee	The primary and highest authority organization for data governance in an organization, responsible for oversight, support, and funding of data governance activities. Consists of a cross-functional group of senior executives. Typically releases funding for data governance and data governance-sponsored activities as recommended by the DGC and CDO. This committee may in turn have oversight from higher-level funding or initiative-based steering committees.
Data Governance Council (DGC)	Manages data governance initiatives (e.g., development of policies or metrics), issues, and escalations. Consists of executive according to the operating model used. See Figure 17.
Data Governance Office (DGO)	Ongoing focus on enterprise-level data definitions and data management standards across all DAMA-DMBOK Knowledge Areas. Consists of coordinating roles that are labelled as *data stewards* or *custodians,* and *data owners*.
Data Stewardship Teams	Communities of interest focused on one or more specific subject-areas or projects, collaborating or consulting with project teams on data definitions and data management standards related to the focus. Consists of business and technical data stewards and data analysts.
Local Data Governance Committee	Large organizations may have divisional or departmental data governance councils working under the auspices of an Enterprise DGC. Smaller organizations should try to avoid such complexity.

Data Governance Operating Model Types
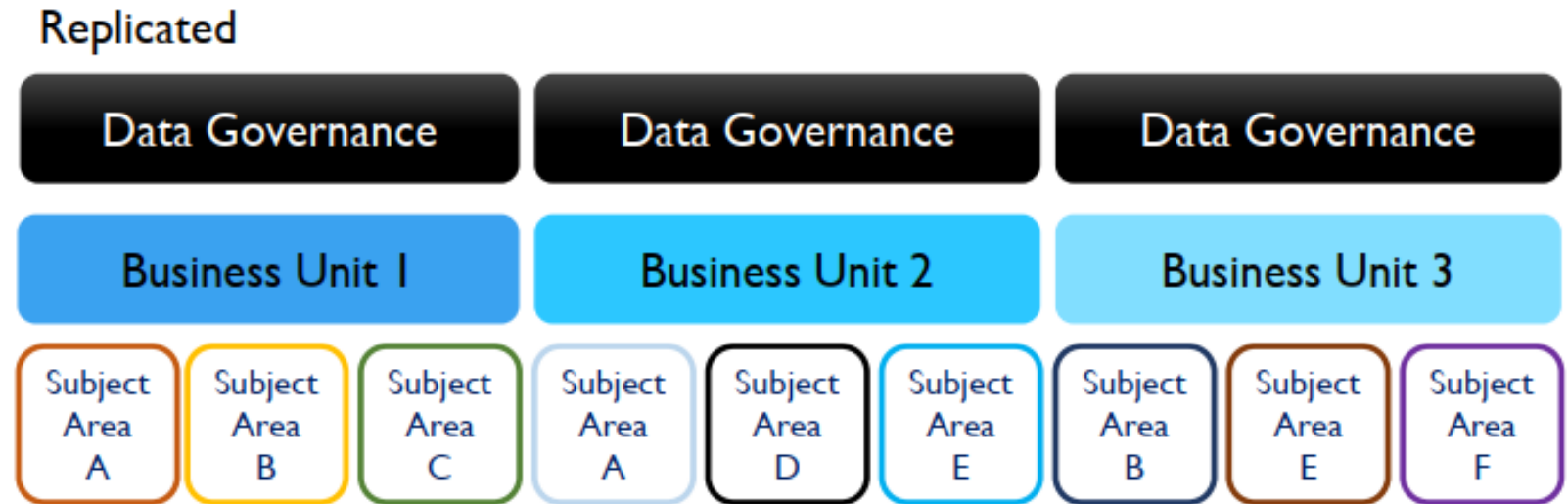
- **Centralized Model**
  - One Data Governance organization oversees all activities in all subject areas
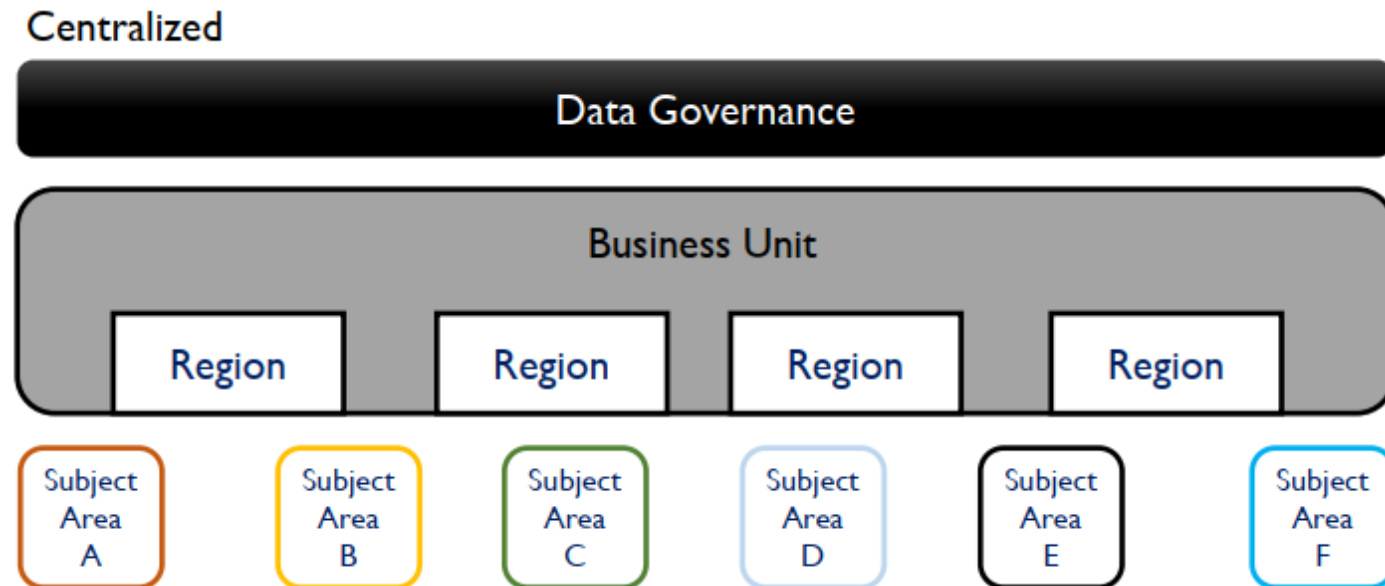
# Data Governance Operating Model Types

- **Replicated Model**
  - The same Data Governance operating model and standards are adopted by each business unit

Replicated

Data Governance	Data Governance	Data Governance
Business Unit 1	Business Unit 2	Business Unit 3

Subject Area A	Subject Area B	Subject Area C	Subject Area A	Subject Area D	Subject Area E	Subject Area B	Subject Area E	Subject Area F

Data Governance Operating Model Types

- **Federated Model**
  - One Data Governance organization coordinates with multiple Business Units to maintain consistent definitions and standards.
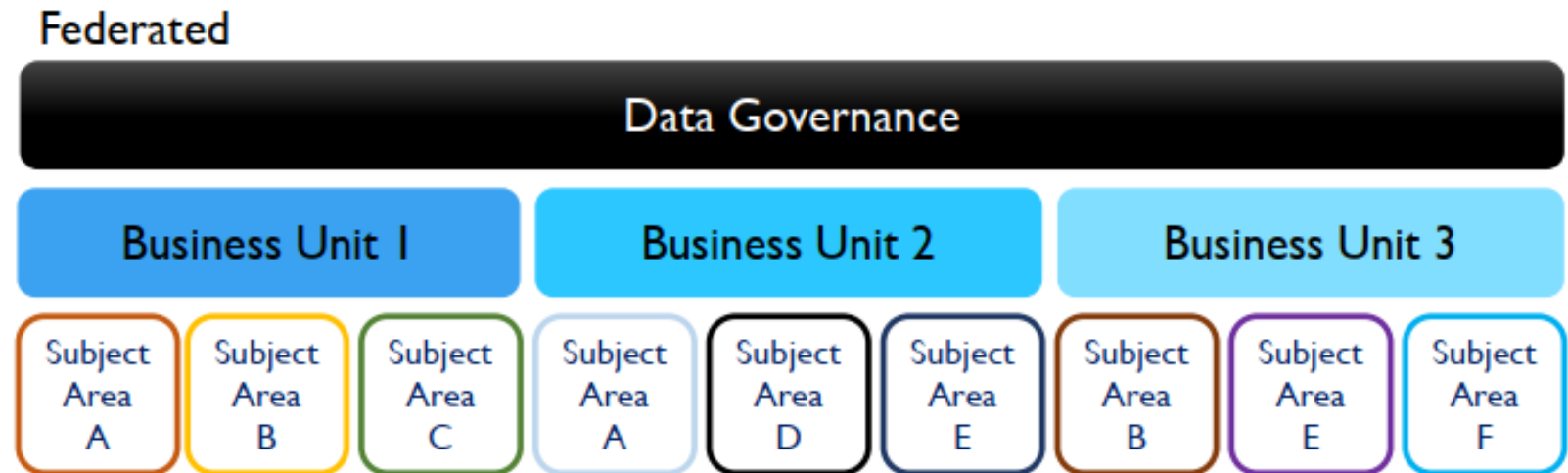
Data Stewardship

- **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
    - can be formalized through job titles and descriptions, or it can be a less formal function
    - Sometimes called: data custodian, data trustee
- Some activities of data stewardship:
    - Creating and managing core Metadata
    - Documenting rules and standards
    - Managing data quality issues
    - Executing operational data governance activities
- **Data stewards** manage data assets on behalf of others (stakeholders) and in the best interests of all the organization

Types of Data Steward

- Formally appointed data stewards may take the following position:
  - **Chief data stewards**: chair of data governance bodies
  - **Executive data stewards**: senior managers, serve a data governance council
  - **Enterprise data stewards**: have oversight of a data domain across business functions
  - **Business data stewards**: business professionals, recognized subject matter experts
  - **Data owners**: business data stewards who have approval authorities for decisions about data within their domain
  - **Technical data stewards**: IT professionals operating within one of knowledge areas, e.g. data integration specialists, database administrators, business intelligence specialists, data quality analysts.
  - **Coordinating data stewards**: lead and represent teams of business and technical data stewards in discussions across teams and with executive data stewards

KNOWLEDGE & SOFTWARE ENGINEERING

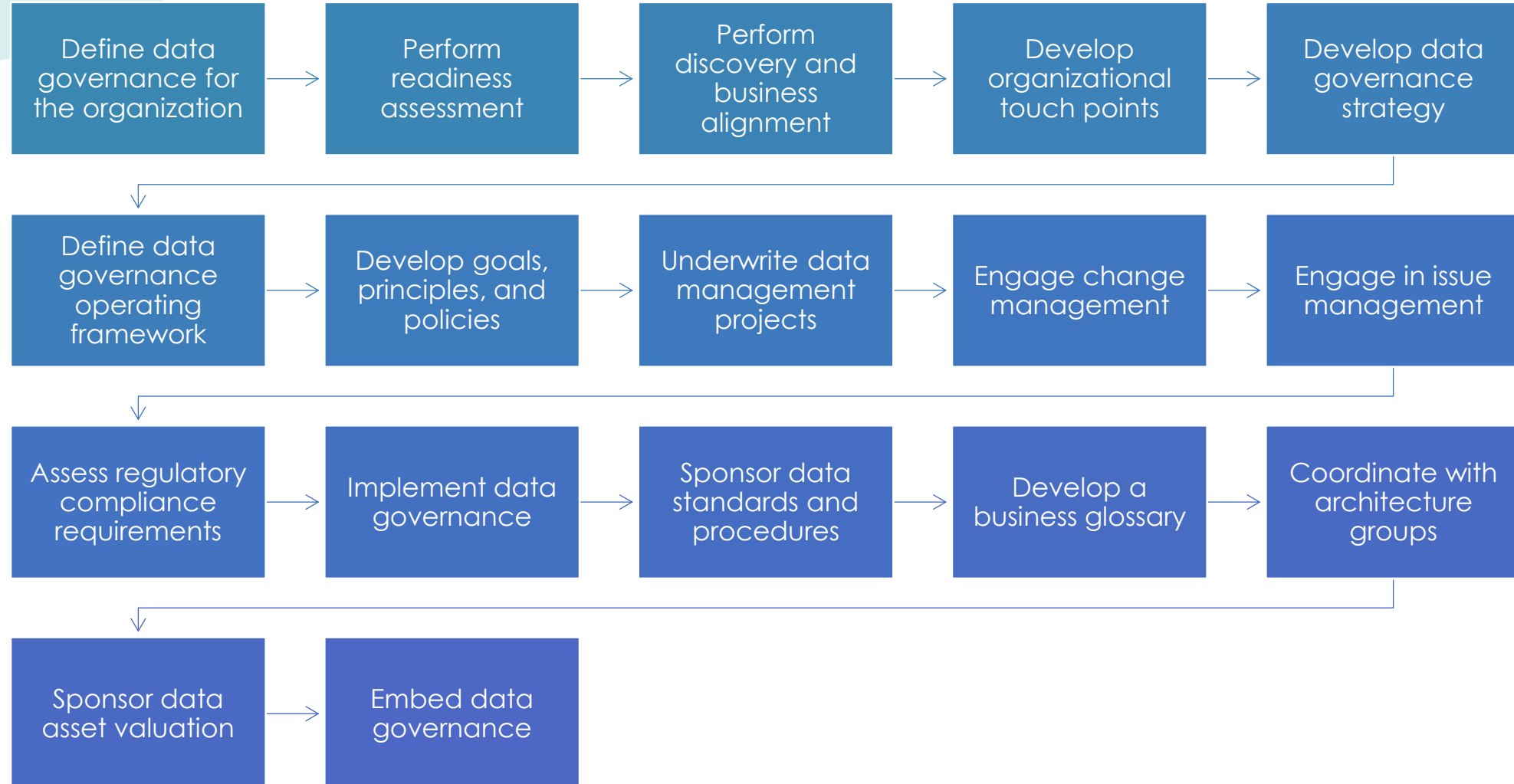# *Other Issues*

- **Data policies**
  - Directives that codify principles and management intent into fundamental rules governing the creation, acquisition, integrity, security, quality, and use of data and information.

- **Data asset valuation**
  - The process of understanding and calculating the economic value of data to an organization.
  - Ways to measure the value of data:
    - Usage of data vs costs of acquiring and storing the data as well as risk mgt.
    - Replacement cost
    - Market value
    - Identified opportunities
    - Selling data
    - Risk cost

KNOWLEDGE & SOFTWARE ENGINEERING

Data Governance Activities

Define data governance for the organization → Perform readiness assessment → Perform discovery and business alignment → Develop organizational touch points → Develop data governance strategy

Define data governance operating framework → Develop goals, principles, and policies → Underwrite data management projects → Engage change management → Engage in issue management

Assess regulatory compliance requirements → Implement data governance → Sponsor data standards and procedures → Develop a business glossary → Coordinate with architecture groups

Sponsor data asset valuation → Embed data governance

KNOWLEDGE & SOFTWARE ENGINEERING

Data Governance Metrics

To measure progress of the rollout of data governance, compliance with the data governance requirements, and the value data governance is bringing to the organization.

Sample metrics:

- Value
  - Contributions to business objectives
  - Reduction of risk
  - Improved efficiency in operations
- Effectiveness
  - Achievement of goals and objectives
  - Extent stewards are using the relevant tools
  - Effectiveness of communication
  - Effectiveness of education/training
  - Speed of change adoption

- Sustainability
  - Performance of policies and processes
  - Conformance to standards and procedures

KNOWLEDGE & SOFTWARE ENGINEERING

Tools and Techniques

- A data governance program must manage its own work and its own data effectively.

- Tools help not only with these tasks, but also with the metrics that support them.

- Some examples of tools and techniques:
  - Online presence/websites
  - Business glossary
  - Workflow tools
  - Document management tools
  - **Data governance scorecards**
  - **Data governance maturity model**

Data Governance Scorecards

- A collection of metrics to track data governance activities and compliance with policies, to be reported up to the data governance council and data governance steering committees
- Key characteristics[1]
  - It is defined at enterprise or business level, usually depending on the operating model
  - It is regularly measured and reported
  - Can be split into individual business data steward scorecards or project level scorecards
- A data governance scorecard can be quite detailed as it includes various categories[1]
  - There is no standard of what these categories should be as it depends on your own implementation, industry, culture, and business needs

Data Governance Scorecards
Example of categories [1)]

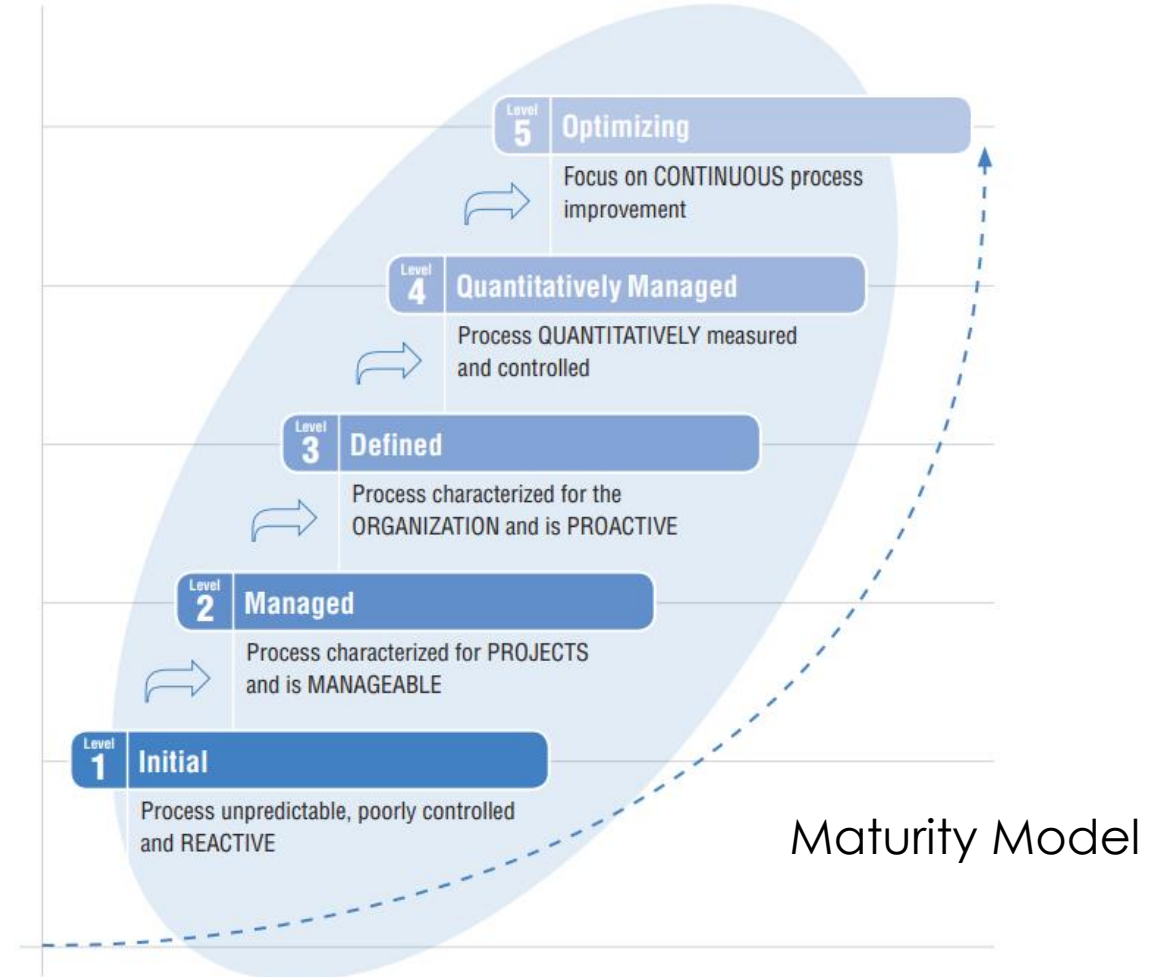| 1. Organization	2. Data infrastructure	3. Data controls	4. Data quality	5. Financial
• *Purpose*: Measures the data skills, training, and overall organizational maturity of its data governance program • *Sample objectives*: Increase data skills, increase data governance awareness & maturity • *Sample KPIs*: # of certified data stewards, employee satisfaction rate	• *Purpose*: Measures the progress on physical data storage, systems and applications • *Sample objectives*: Reduce legacy databases, improve master data usage • *Sample KPIs*: # of systems using master data, % reduction in total cost of ownership	• *Purpose*: Measures the organization's compliance over its data standards, including the completeness of its metadata • *Sample objectives*: Automate data controls and audits, increase compliance to data standards • *Sample KPIs*: Compliance test results, % decrease in policy failure	• *Purpose*: Measures improvements in data quality dimensions • *Sample objectives*: Improve data quality dimensions • *Sample KPIs*: % improvement in data completeness, data accuracy rate	• *Purpose*: Tracks the return of investment of data governance at project and program level • *Sample objectives*: Ensure costs are in line with budget • *Sample KPIs*: Total budget to actuals, % of operating budget savings

Data Governance Maturity Model

- A tool that helps organizations assess the current state of their data governance program [2]

- Two models to be discussed in brief:
    - IBM Data Governance Maturity Model [3]
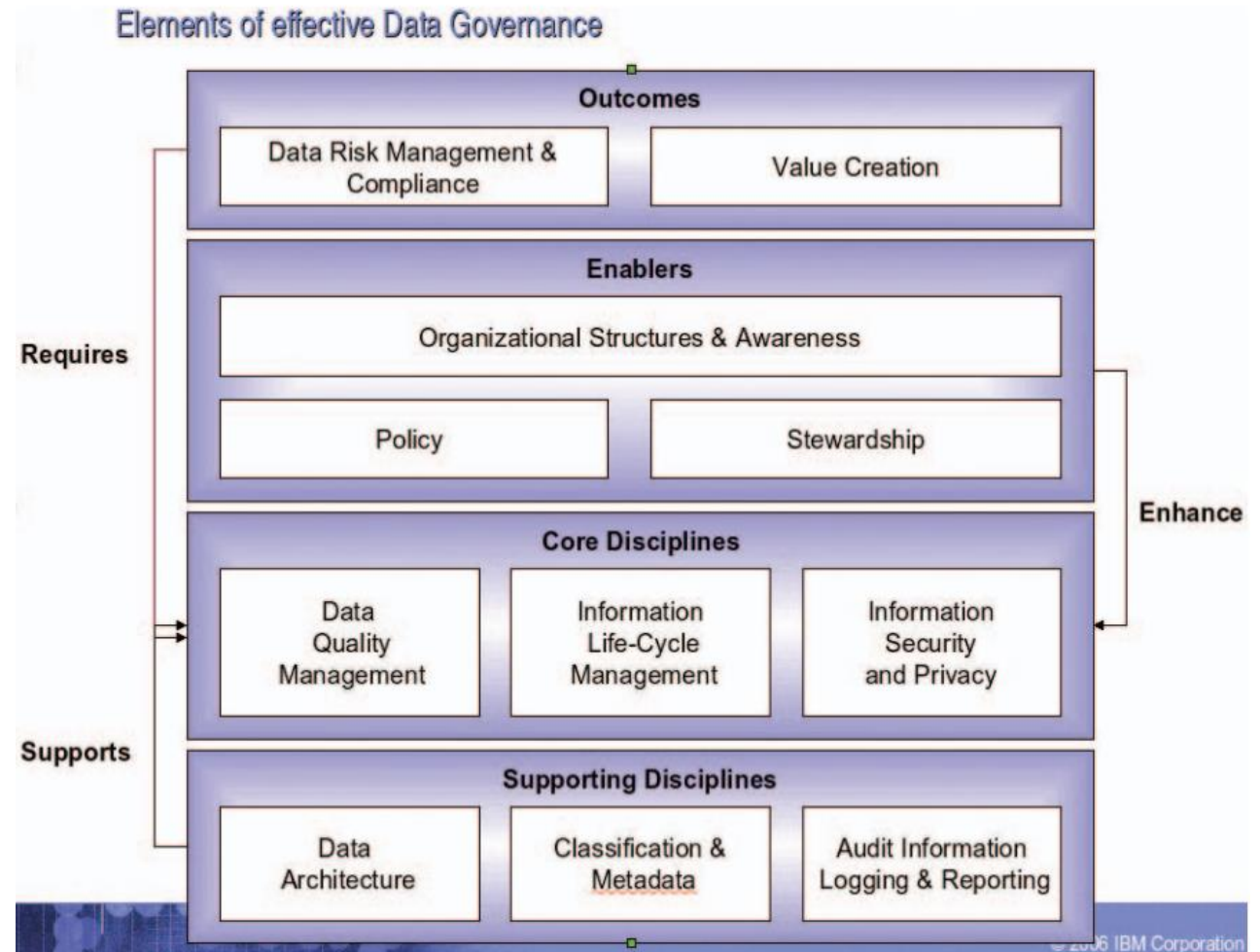   - Stanford Data Governance Maturity Model [4]

IBM Data Governance Maturity Model [3)]

- Published in October 2007
- Based on CMM (Capability Maturity Model)
  - a methodology used to develop and refine an organization's software development process, developed by Software Engineering Institute (SEI) in 1984


Level 5 Optimizing
Focus on CONTINUOUS process improvement

Level 4 Quantitatively Managed
Process QUANTITATIVELY measured and controlled

Level 3 Defined
Process characterized for the ORGANIZATION and is PROACTIVE

Level 2 Managed
Process characterized for PROJECTS and is MANAGEABLE

Level 1 Initial
Process unpredictable, poorly controlled and REACTIVE

Maturity Model

IBM Data Governance Maturity Model[3]

- The model helps assess and measure progress within each of the 11 data governance domains

Elements of effective Data Governance

Stanford Data Governance Maturity Model [4]

- Developed in 2011 by Stanford University's Data Governance Office

- The model was adapted from other models, such as IBM's and CMM's

- It is based on the structure of their data governance program, with a focus on both foundational and project aspects of data governance

The Stanford DG Maturity Model
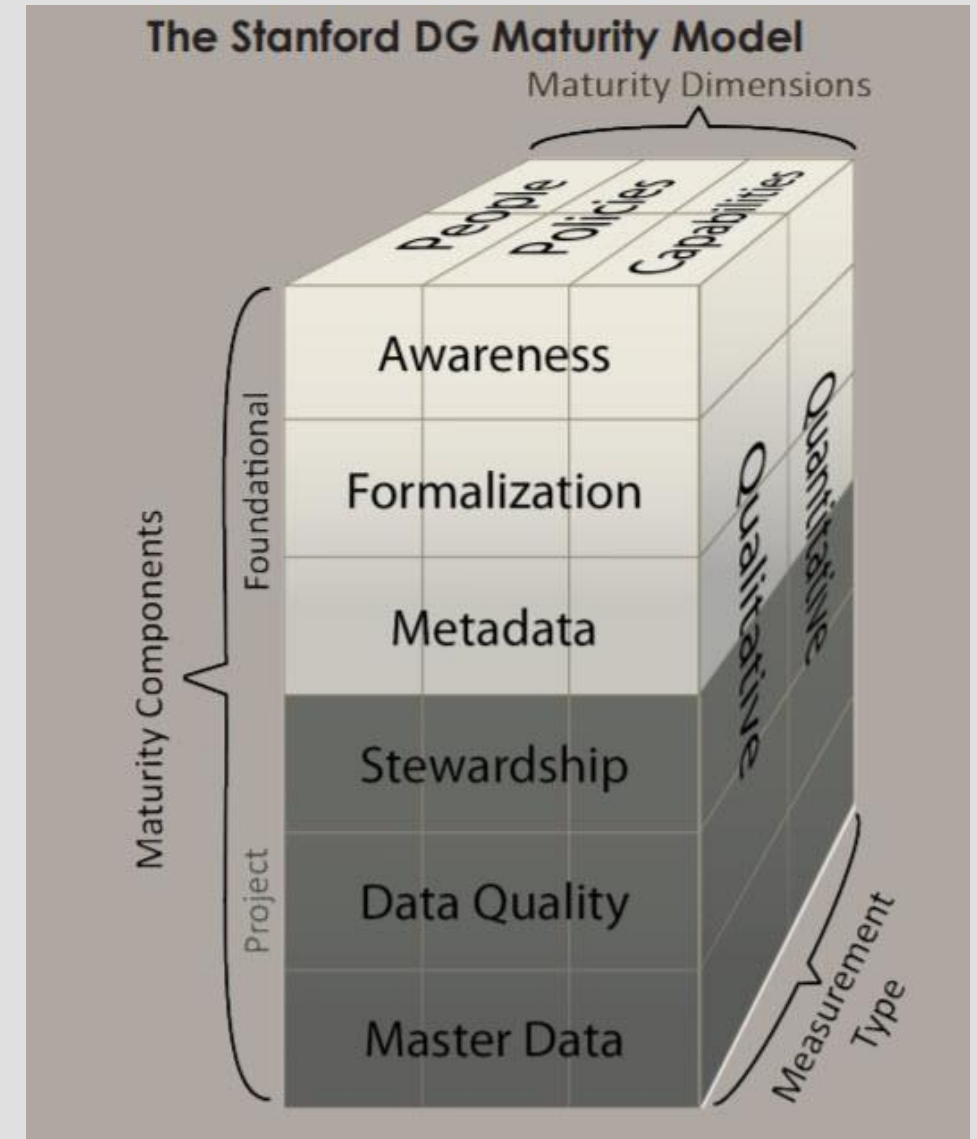
Stanford Data Governance Maturity Model 4)

Data Governance Maturity Model			
Guiding Questions for each Component-Dimension			
Foundational	People	Policies	Capabilities
Awareness	What awareness do people have about the their role within the data governance program?	What awareness is there of data governance policies, standards and best practices?	What awareness is there of data governance enabling capabilities that have been purchased or developed?
Formalization	How developed is the data governance organization and which roles are filled to support data governance activities?	To what degree are data governance policies formally defined, implemented and enforced?	How developed is the toolset that supports data governance activities and how consistently is that toolset utilized?
Metadata	What level of cross functional participation is there in the development and maintenance of metadata?	To what degree are metadata creation and maintenance policies formally defined, implemented and enforced?	What capabilities are in place to actively manage metadata at various levels of maturity?

Project	People	Policies	Capabilities
Stewardship	To what degree have stewardship roles been defined and filled?	To what degree are stewardship policies defined, implemented and enforced?	What capabilities are implemented to support the effective stewardship?
Data Quality	To what degrees have data quality competencies developed?	To what degree are data quality policies defined, implemented and enforced?	What capabilities are implemented to support the production and maintenance of high quality data?
Master Data	To what degree has a formal master data management organization been developed and assigned consistent responsibilities across data domains?	To what degree are master data policies defined, implemented and enforced?	What capabilities are available and implemented to actively master and provision master data?

Online Sources

1) G. Firican, "How to create a data governance scorecard – Free template", https://www.lightsondata.com/data-governance-scorecard-free-template/

2) "Data Governance Maturity Model: A Roadmap to Optimizing Your Data Initiatives and Driving Business Value", https://atlan.com/data-governance-maturity-model/#:~:text=A%20data%20governance%20maturity%20model%20is%20a%20tool%20that%20helps,against%20those%20of%20similar%20organizations.

3) "The IBM Data Governance Council Maturity Model: Building a roadmap for effective data governance, http://www.databaser.net/moniwiki/pds/DataWarehouse/leverage_wp_data_gov_council_maturity_model.pdf

4) G. Firican, "Stanford data governance maturity model", https://www.lightsondata.com/data-governance-maturity-models-stanford/