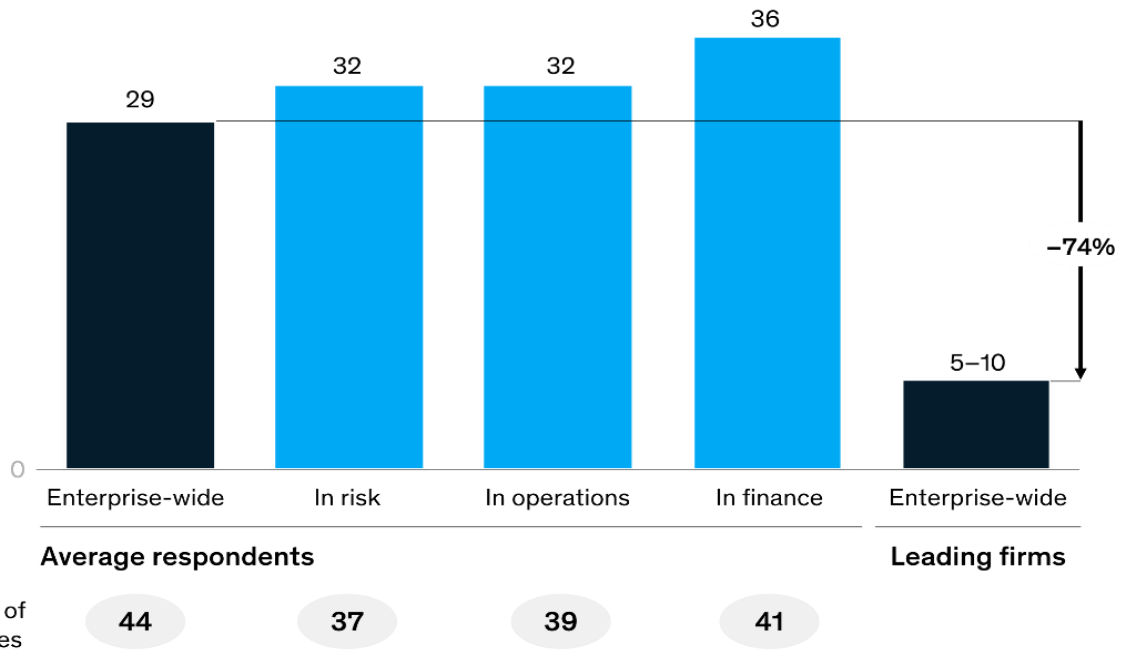
Case Studies

IS465: Data Management and Governance

Designing data governance that delivers value

Lack of data quality and availability can cause employees to spend a significant amount of time on non-value-added tasks.

Time spent on non-value-added tasks due to poor data quality and availability[1]
Estimated % of total employee time

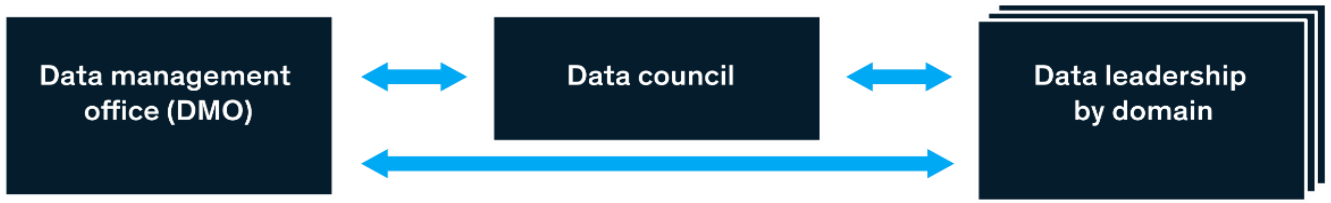


[1] Data sourcing, data aggregation, data reconciliation, data cleansing, manual reporting, etc.
Source: McKinsey Global Data Transformation Survey, 2019

McKinsey & Company

Designing data governance that delivers value

A best-practice data-governance model typically includes three organizational components.



- Defines policies and standards
- Empowers data leaders in the domains with tools, playbooks, training, etc
- Ensures coordination and consistency across key roles in the data life cycle
- Links data-governance efforts to technology and other related functions
- Provides targeted support for issues

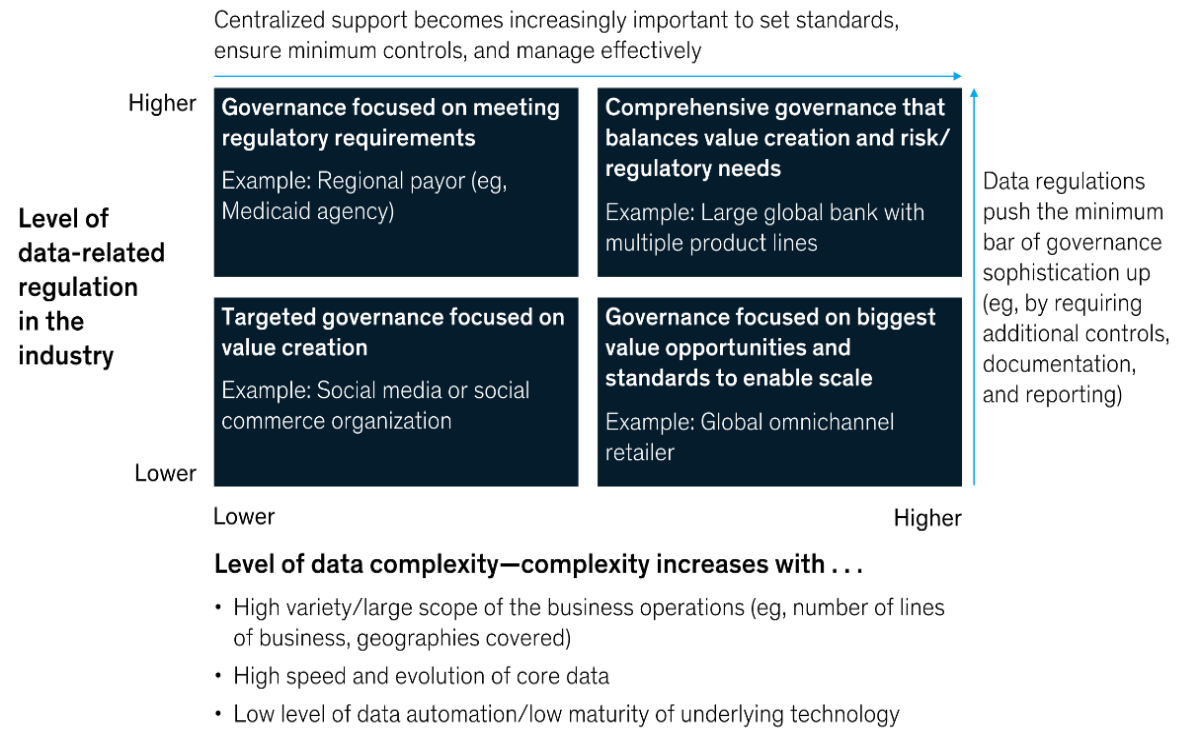
- Owns strategic direction and principles
- Proposes DMO structure, defines domains, and assigns leaders
- Confirms adherence to standards and policies
- Reviews progress on initiatives and serves as funding gateway
- Resolves and/or escalates issues
- Chaired by head of DMO

- Sets and executes strategy of the domain—initiatives, external sources, etc
- Understands and meets needs of data consumers
- Owns and manages the data—assesses and improves data quality, defines definitions and data model, etc
- Participates in the data council

McKinsey & Company

Designing data governance that delivers value

Data-governance archetypes can be used to inform the level of sophistication needed.



McKinsey & Company

Uber set up a flexible approach to data governance

- Uber operates in 70+ countries, each with its own data governance framework to comply with local and regional laws.

- To ensure compliance across all regions,
  - Uses a core platform to take care of data privacy and security centrally.
  - Collects data globally but adapts its governance policies depending on the origins of each dataset by using customizations and plugins.

Uber set up a flexible approach to data governance

- Manikandan Thangarathnam, the Senior Director of Mobility and Platforms at Uber, puts it:
  - *"For example, a city or a country may have a rule by which they do not want to expose the driver's name to the rider. So we have the capability to switch on or off depending on the city; we do not have to create a different application for that. That is the amount of engineering and flexibility that we have built into our system."*
- The company has invested in training its people and honing their skills to ensure the success of its data governance programs across all regions.

How Wells Fargo built a single source of truth to ensure proper governance

- Wells Fargo's data governance strategy highlights the importance of creating a single source of truth to enhance data accuracy and reliability.
- Their approach was to centralize data from multiple sources to create a unified, trustworthy source that reduced discrepancies and improved consistency.
- This streamlined data management and allowed for more accurate reporting and analysis, thereby enhancing decision-making across the organization.

How Wells Fargo built a single source of truth to ensure proper governance

- Prahalad Thota, the former Senior Vice President, Head of Enterprise Analytics & Data Science at Wells Fargo, describes their efforts:
  - *"We established a team that brings data together and organizes it into our Enterprise Data League. The team focuses on establishing the right governance, in terms of storage and management, for the data. The team also focuses on leveraging data for driving cutting-edge use cases."*
- Wells Fargo's data governance strategy also emphasized the importance of data visualization (using Tableau) to make data more accessible to non-technical stakeholders and improve overall data literacy.
- As a result, Wells Fargo was able to ensure data consistency, accuracy, and visualization to reduce risks and foster data-driven decision-making.

Benchmarking Data Governance Maturity

- Project Background
  - A construction management firm was seeking:
- An assessment of their data governance needs and advisory support for selecting a tool to address them.
- To clearly define and prioritize their data management and governance use cases.

Data Governance Needs Assessment & Use Case Definition

Tool Evaluation Matrix & Vendor Recommendation

Data Governance Solutions Architecture Recommendation

Data Governance Findings

- Categorize key takeaways into five data governance themes.

People

- Focus on standardizing data governance processes
- Create an organization where individuals are eager to engage in potential data management initiatives
- Encourage staff to upskill and pursue data literacy training to help fill data governance gaps in their daily operations

Process

- Identify gaps in standardized processes for data governance
- Document and communicate processes across the organization, and define roles and responsibilities

Content

- Identify underutilized customer and project data
- Correlate previously unrelated data to generate new insights
- Ability to gain insights from diverse types of data (structured/ semi-structured/ unstructured)

Technology

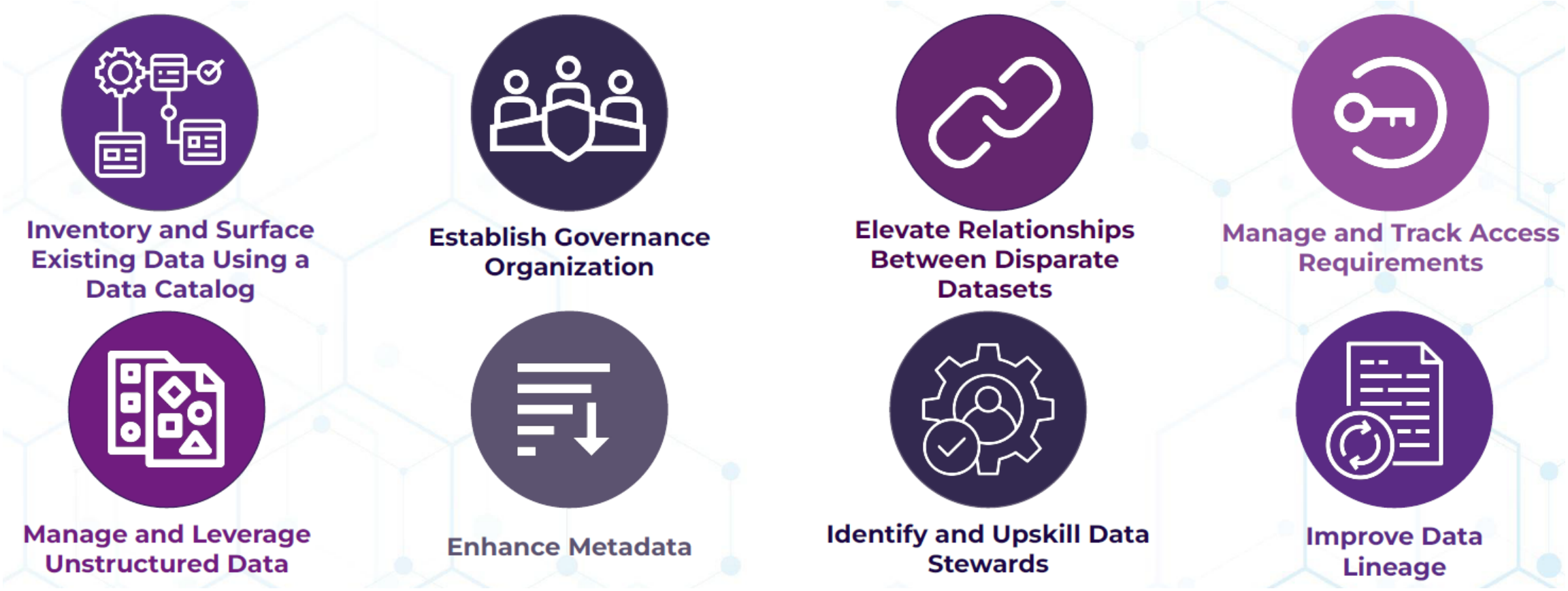
- Identify potential tools for better unifying and governing their data
- Unify data across source systems for an enterprise-level view regardless of physical location of data
- Build more automated integrations for business opportunities, projects, and skills

Culture

- Create the support and engagement to pursue initiatives for overarching data governance and management needs
- Prioritize a data-centric culture that enables discovery and connection of data

Priorities to Address Data Governance Challenges

- A selection of recommendations.



End User Training

- Constructed a comprehensive training plan to upskill the organization in five key areas.

People	Process	Content	Technology	Culture
• Training focused around identifying the correct structure for creating a data stewardship hierarchy.	• Training focused on how to establish data governance actions around Usage, Access, and Sharing at an enterprise organization.	• Training providing introductory guidance on the concepts of data governance and data stewardship.	• Training providing an introduction to business stakeholders on the concept of a data catalog and the value it will bring to their divisions.	• Training focused on how to communicate the value of data stewardship to various business stakeholders within an organization.

Data Stewardship Hierarchy

Stewardship Council: Decision-making body of the data stewardship team.

Stewardship Lead: Member of the data stewardship council.

System Administrators: Member of the data stewardship council.

Strategic Leads: Member of the data stewardship council.

Data Stewards: Do not have voting rights, but inform the stewardship council during decision-making processes.

Provides strategic direction and collective decision-making.

Guides meetings and draws on stewardship best practices to inform decisions. Manages requests.

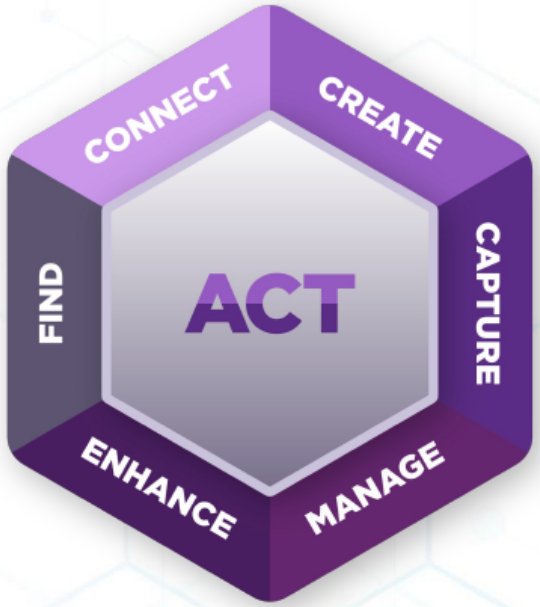
Offers a technical perspective around suggested stewardship changes and leads changes in systems.

Guides data stewardship evolutions by providing insight into business and strategic objectives.

Offers program area or operational area perspective on stewardship needs.

Each recommended role builds on the preceding one in terms of responsibility level and decision power.

Access, Usage, and Sharing



CONNECT	**Aggregate** User(s) can create a dataset that is made up of but not linked to existing data assets.	**Branch** User(s) can create a dataset that is made up of and linked to existing data assets.
CREATE	**Draft** User(s) can add a new data asset.	
MANAGE	**Manage/Own** User(s) can archive or delete a data asset. **Comment** User(s) can view a data asset and attach non-edit notes.	**Govern** User(s) can determine the standards and structure of a data system.
ENHANCE	**Edit** User(s) can add to, modify, or remove pieces of an existing data asset.	
FIND	**See** User(s) can see the data asset in a list of results, but cannot view it.	**View** User(s) can read the contents of a data asset.

Tools and Frameworks

- A multi-national financial organization was facing difficulties in unifying governance, discovery, and search across multiple metadata storage platforms within their global enterprise.
- In order to rectify their existing data quality and governance issues with a standardized metadata platform, this organization identified a data catalog as a foundational solution to start addressing these challenges and consulted with us to lead the implementation.

Increased data maturity and development

8,000+
Business glossary objects

A new metadata platform that enabled more advanced use cases

5,500+
resources

Maintain compliance tracking through external ontologies

Why is a Data Catalog Foundational?

- A data catalog serves as a data governance tool that allows us to collect, aggregate, and present logical and physical metadata to end users.

- A modern data catalog….
  - Contextualizes and enriches information with meaning of data based on business or data domains.
  - Establishes relationships across disparate data sources and across business and technical concepts.
  - Unifies unstructured and structured data to connect data of all formats.
  - Makes data and information easily searchable and discoverable.

Data Catalog Business Value

COST SAVINGS & INCREASED REVENUE

Provide Structure	• Tag data using terms from a customized business glossary • Increase the accuracy and range of search
Improve Findability	• Target access to data to specific audiences • Enable faster access to the right data and the people who manage it
Improve Discoverability	• Implement a user-centric and scalable data inventory • Help users organize, find, and discover data
Reuse Content	• Find and connect existing data for reuse • Minimize duplication of existing data and dashboards • Standardize data schemas across sources
Integrate Sources	• Integrate multiple disparate sources of data • Connect both structured and unstructured datasets

How a Data Catalog Fits into Data Governance Tools

A **data fabric** enables data federation and virtualization of semantic labels or rules (e.g. taxonomies/ business glossaries or ontologies) to capture and connect data based on business or domain meaning and value.