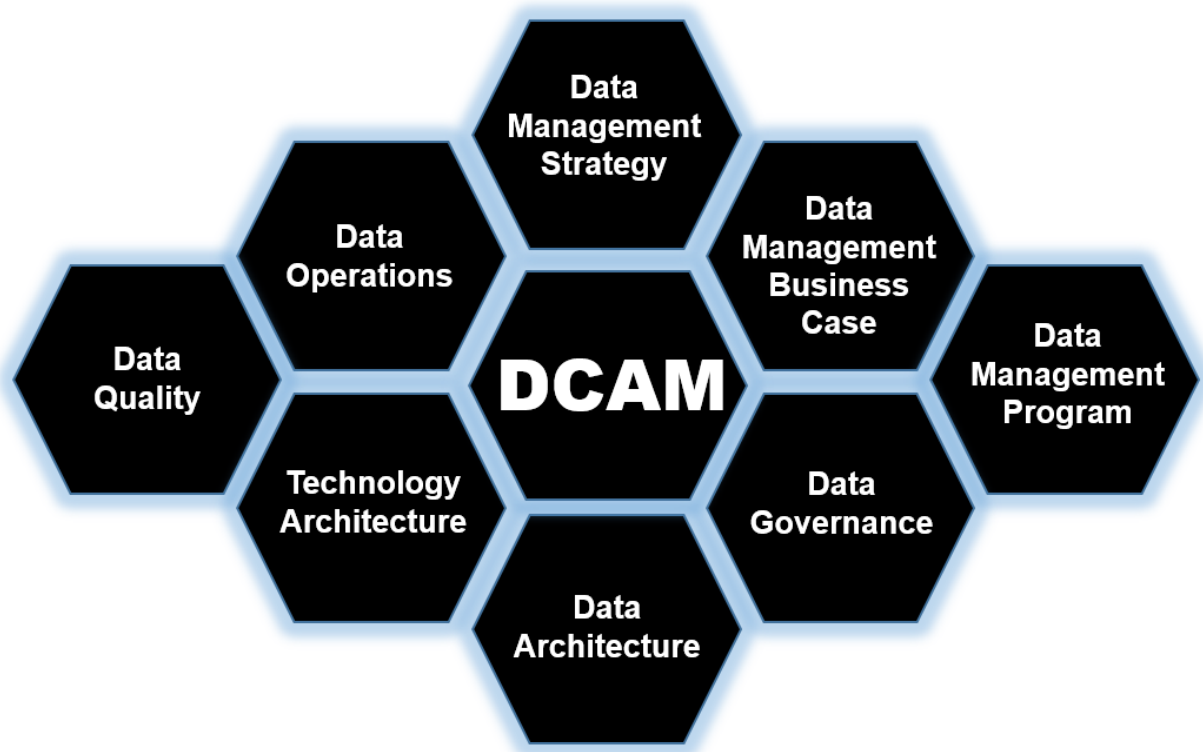# Data Management
Capability Assessment Model
(DCAM)

Working Draft Version 0.7
July 30, 2014

Developed by

Published: July 30, 2014
EDM Council
WORKING DRAFT – Version 0.7

INTRODUCTION

The Data Management Capability Model (DCAM) was created by the Enterprise Data Management Council based on the practical experiences and hard won lessons of many of the world's leading organizations. It is a synthesis of best practices associated with the management of data content across the horizon of interconnected processes. The Data Management Capability Model defines the scope of capabilities required to establish, enable and sustain a mature data management discipline. It addresses the strategies, organizational structures, technology and operational best practices needed to successfully drive data management. It addresses the tenets of data management based on an understanding of business value combined with the reality of operational implementation.

To manage data in today's organizational environment starts by recognizing that proper data management is about managing data as "meaning". This is a relatively new concept for many organizations. It is not easy to articulate and not very well understood. Data exists everywhere within an organization and must be managed consistently within a well-defined control framework. The DCAM helps identify this framework by defining the capabilities required to make data management a critical part of a firms' everyday operational fabric.

The challenges of properly managing data are significant. There are many legacy repositories and a plethora of functions to unravel. There are social and political barriers to overcome. There are real IT challenges and execution gaps to address. Data ownership and accountability are hard to implement. Funding is often project based. And many firms simply don't have the strong executive support that is needed to ensure that the organization stays the course in the face of short term measurement criteria, operational disruption and conflicting stakeholder challenges to properly address the realities of the data management challenge.

We understand this reality because we've been there and we have the scars across our back to prove it. Data is foundational. It is the lifeblood of the organization. The "bad data" tax is a significant expenditure for many firms. Unraveling silos and harmonizing data is the prerequisite for eliminating redundancy, reducing reconciliation and automating business processes. Managing data is essential if we are to gain insight from analytics, feed our models with confidence, enhance our service to clients and capitalize on new (but often fleeting) business opportunities. DCAM provides the guidance needed to assess current state, and provide the objectives of target state, for your data program.

The DCAM is organized into eight core components.

1. The **Data Management Strategy** discusses the elements of a sound data strategy, why it is important and how the organization needs to be organized to implement.

2. The **Data Management Business Case and Funding Model** addresses the creation of the business case, its accompanying funding model and the importance of engaging senior executives and key stakeholders for approval.

3. The **Data Management Program** discusses what's organizationally needed to stand up a sustainable Data Management Program.

4. **Data Governance** defines the operating model and the importance of policies, procedures and standards as the mechanism for alignment among (and compliance by) stakeholders.

5. **Data Architecture** focuses on the core concepts of "data meaning" – how data is defined,

described and related.

6. ***Technology Architecture*** focuses on the relationship of data with the physical IT infrastructure needed for operational deployment.

7. ***Data Quality*** refers to the concept of fit-for-purpose data and the processes associated with the establishment of both data control and data supply chain management.

8. ***Data Operations*** defines the data lifecycle process and how data content management is integrated into the overall organizational ecosystem.

Each component is preceded with a definition of what it is, why it is important and how it relates to the overall data management process.   These are written for business and operational executives so as to demystify the data management process.   The components are structured into 35 capabilities and 109 sub-capabilities. These capabilities and sub-capabilities are the essence of the DCAM. They define the goals of data management at a practical level and establish the operational requirements that are needed for sustainable data management.   And finally, each sub-capability has an associated set of measurement criteria to be used in the evaluation of your data management journey.

Welcome to the world of data management.   The EDM Council is indebted to the dozens of members who have contributed to the development of the Data Management Capability Model. We are always searching for ways to enhance and improve the model. We encourage your feedback. We are interested in your rants, raves and alternative points of view. For more information on the DCAM and on the EDM Council, please contact us at info@edmcouncil.org

Michael Atkin   John Bottega
Managing Director   Executive Advisor
EDM Council   EDM Council
atkin@edmcouncil.org   jbottega@edmcouncil.org

Contents

FOREWARD

The concept of data as a foundational component of business operations has arrived. It is now understood as one of the core factors of input into the full spectrum of business and organizational processes. The common theme for firms that are effective in their use of data to reduce operational costs, automate manual processes, consolidate redundant systems, minimize reconciliation and enhance business opportunities is the implementation and management of a **data control environment**. The reason why firms implement a control environment is to ensure trust and confidence among consumers that the data they are relying on for business processing and decision-making is precisely what they expect it to be – without the need for manual reconciliation or without reliance on data transformation processes.

The core components associated with the implementation of a control environment are needed to ensure that all data elements/attributes are precisely defined, aligned to meaning, described as metadata and managed across the full data lifecycle. The key to establishing a control environment however, is the achievement of "unambiguous shared meaning" across the enterprise as well as the governance of the processes related to ensuring definitional precision. Data must be consistently defined because it represents a real thing (i.e. a product, client, account, counterparty, transaction, legal entity, location, process, etc.). All other processes are built upon this foundation.

In a fragmented data environment (the opposite of a control environment) applications development can result in ad hoc naming conventions which exacerbate the problem of common terms that have different meanings, common meanings that use different terms and vague definitions that don't capture critical nuances. For many firms this challenge can be debilitating because there are thousands of data attributes, delivered by hundreds of internal and external sources, all stored in dozens of unconnected databases. This fragmentation results in a continual challenge of mapping, cross-referencing and manual reconciliation. In order to achieve a control environment, every data attribute must be understood at its "atomic level" (as a fact) that is aligned to business meaning without duplication or ambiguity. Managing data as meaning is the key to alignment of data repositories, harmonization of business glossaries and ensuring that applications dictionaries are comparable.

Achieving alignment on business meaning (including the process of how terms are created and maintained) can be a daunting task. It is not uncommon to experience resistance from business users and IT - particularly when there are multiple existing systems linked to critical business applications. The best strategy for reconciliation in a fragmented environment is to harmonize on the legal, contractual or business meaning rather trying to get every system to adopt the same naming convention. Nomenclature represents the structure of data and unraveling data structures/data models are expensive and not necessary. It is better to focus on precisely defining business concepts, documenting transformation processes and capturing real-world data relationships. Once established, existing systems, glossaries, dictionaries, repositories, etc. can be cross-referenced to common meaning.

Managing data as meaning is the cornerstone of effective data management. It needs to be managed along with other "metadata" to ensure consistency and comparability across the enterprise. The other components of metadata can be organized into three core categories: descriptive metadata (i.e. information that identifies where data is located); structural metadata (i.e. information about the physical data layer and how the data is structured) and administrative metadata (i.e. information about when the data was created, its purpose and access rights). Data meaning and metadata management are best understood as the core of your content infrastructure and the baseline for process automation, applications integration and alignment across linked processes.

The implementation and management of a control environment is governed by standards, policies and

procedures. These are the essential mechanisms for establishing a sustainable data management program and for ensuring compliance with a control environment in the face of organizational complexity. Managing meaning is the key to effective data management. Meaning is achieved through the adoption of semantic standards. Standards are governed by policy. Policy is established by executive management, supported by data owners and enforced by Corporate Audit. Get the data infrastructure established and governed – it represents the foundation for operational efficiency and must not be compromised.

Control Environment Capability Objectives

1. The concept of a control environment is understood by relevant stakeholders and adopted by the organization (standards-based, harmonized across lifecycle, unique identifiers, aligned to meaning). The organization recognizes the need for a control environment to meet business, operational and regulatory objectives (*see data management strategy*)

2. The components associated with a control environment have been defined, verified by stakeholders (i.e. inventoried and confirmed), aligned with technical capability and approved by executive management.   Policies, procedures and standards exist for all relevant areas including data quality, data access/distribution, authorized use/entitlement control, data privacy and data security

3. The framework for implementing a control environment, including reconciliation of disparate systems, have been fully resourced (*see business case and funding*)

4. The standards that are needed to implement the control environment are defined and verified by stakeholders (for relevant products, accounts, clients, business partners, legal entities, counterparties, vendors, etc.). Business processes are identified, documented and aligned with data requirements.

5. Data attributes for relevant business processes are known, segmented according to criticality and understood in the context of how data is compounded/how derived data elements are calculated. Existing systems, processes, repositories and consuming applications across the full data lifecycle are mapped to the control environment standards and aligned with systems of record. Rules and conversion procedures for transformation and cross-referencing are documented. Shared data attributes are identified and mapped to processes and sub-processes.

6. Standard identifiers, metadata and taxonomies are established and integrated across the enterprise for all functions and processes. The process for new standards adoption is documented and implemented

7. Data in all repositories are aligned to "common meaning" as an ontology. The ontology is modeled and verified by SMEs. There is a common method for defining, achieving agreement, updating and promulgating the concept of "single term, single definition" based on how business processes work in the real world. All changes to the corporate ontology are synchronized and aligned to the systems of record

8. Procedures are in place to manage changes and exceptions to the control environment

9. A centralized/aligned metadata repository is implemented and maintained. The metadata repository is managed from descriptive, structural and administrative dimensions

10. All new product development initiatives, data integration activities and data consolidation efforts use the control environment standards

11. The control environment is governed across the enterprise with clear accountability. The governance process consists of a combination of IT infrastructure, program management offices, data administrators and data owners (*see governance*)

12. Compliance with the control environment is monitored, measured and audited. Results of the compliance audit is shared with executive management

13. Communications mechanisms are in place to ensure that the goals, policies and procedures of the control environment are implemented; that business and IT can communicate with each other; that issues can be escalated as appropriate; that priorities are established; that policies and standards are implemented and that employees are in compliance with the control processes (*see governance*)

14. There is close cooperation between the Board of Directors, executive management, lines of business, information technology and operations on the implementation and management of the control environment. Stakeholders receive training in the policies that exist and the procedures that need to be followed to achieve organizational compliance

15. Executive management sets expectations and gives authority to implement the control environment. Expectations are translated into incentives and operational constraints. Lines of business are managing within established boundaries. Performance is linked to implementation of the control environment
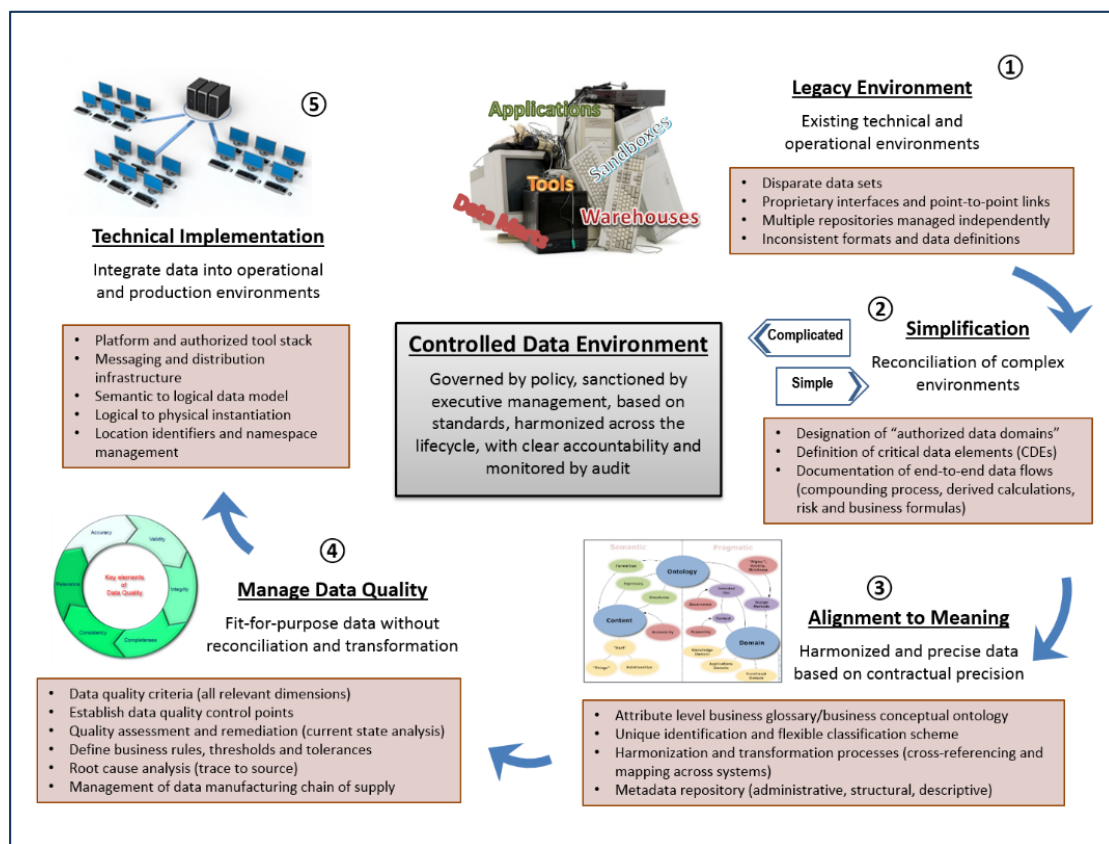
Scope of work required to achieve a data control environment

1.0   DATA MANAGEMENT STRATEGY

Definition:
The Data Management Strategy determines how data management is defined, organized, funded, governed and embedded into the operations of the organization. It defines the long-term vision including a description of critical stakeholder or stakeholder functions that must be aligned. Data Management Strategy demonstrates the business value that the program will seek to achieve. It becomes the blueprint (or 'master plan') that describes how the organization will evaluate, define, plan, measure and execute a successful and mature data management program.

Purpose:
The purpose of developing a Data Management Strategy is to articulate the rational for the data management program. The strategy defines "why" the program is needed as well as the expected benefits, goals and objectives. The strategy also describes "how" to align and mobilize the organization in order to implement a successful data management program.

Introduction:
A Data Management Strategy defines the organizational rationale for implementing a data management program, explains what the overall program aims to achieve and identifies how the various components of the initiative fit together.   A strategy accurately reflects the requirements of the data consumers in order to give confidence to stakeholders that the data management program will be valuable, practical and managed in an effective manner. A data management strategy should emphasize the importance of collaboration as well as the data challenges that result due to the interconnected nature of business processes.

A data management strategy defines the overall framework of the program. It should be structured to address the core principles of data management so that critical stakeholders can understand the value of a data management program as it relates to their functions and strategic initiatives.

A Data Management Strategy needs to:
- Articulate the scope of the data management program
- Establish the priorities for phased implementation
- Provide the guidance for establishing the data governance framework
- Express the importance of developing a data quality program
- Reinforce the use of data content standards.
- Reflect practical implementation reality and alignment to IT and operational capabilities
- Define rational timeframes for implementation
- Address the importance of establishing and staffing the data management program function
- Address the importance of developing a sustainable funding model
- Address the importance of developing evaluative criteria to measure and monitor program progress and effectiveness.

Central to a data management strategy is the articulation of the "target state". An effective data management strategy describes target state objectives, identifies key stakeholders, discusses organizational structure, accountability and describes the need for discipline and governance. More importantly, a strategy identifies operational inefficiencies and gaps. It is important for a strategy to compare target state to current state in order to show the organizational, functional and technological gaps, and then demonstrate the how these gaps will be closed.

Data management strategy is not static and must be able to evolve as the need of the organization change. The most effective and successful data management strategies are those that are visibly endorsed by executive management and are supported by mandatory organizational policy.

Goals:

→ Define a strategy that is aligned with the goals and objectives of the organization and ensure this strategy is approved by all relevant business, technology, operational and executive stakeholders

→ Explain the importance of establishing a recognized and sustainable data management "program". Define the need for metrics to assess the program and to ensure alignment with established cost/benefit evaluation methodologies.

→ Capture high level data requirements. Ensure all relevant corporate audit and regulatory issues have been identified and that key stakeholders understand and agree to the high level requirements.

→ Define the process for determining scope and priorities of the Data Management Program. Ensure that the scope of the Program is aligned with defined business value and organizational priorities.

→ Make sure the Program can be practically implemented from both a technical and architectural perspective.

→ Identify high-level immediate; transitional and long-term deliverables as well as associated resource and funding requirements necessary to implement and sustain the data management program.

→ Ensure the data management strategy is clearly articulated and communicated across the organization and is reflected in architectural technology planning.

Core Questions:

→ Does the Data Management Strategy clearly articulate the reason and the importance of implementing a separate Enterprise Data Management Program?

→ Is there executive, operational, technology and business buy-in? Do stakeholders agree to support and sustain such a program?

→ Has the Data Management Strategy sufficiently framed the immediate, medium and long-term goals and objectives of the Data Management Program in line with organizational priorities?

→ Has the Data Management Strategy effectively identified the critical areas of focus including how priorities are established and verified?

→ Has the Data Management Strategy identified staffing resources, operating model and the funding approach needed to establish, lead and maintain the Data Management Program?

Capabilities and Assessment Criteria:

1. DATA MANAGEMENT STRATEGY

1.1. A Data Management Strategy (DMS) is Specified and Shared with Relevant Stakeholders

1.1.1. DMS is developed

The data management strategy needs to be documented in collaboration with the full spectrum of business, technology and operations management.

Capability Objectives

- DMS has been documented.
- DMS has been aligned with business, technology and operations.
- DMS has been published to all relevant stakeholders.

1.1.2. The DMS is aligned with the high-level organizational objectives.

High level organizational objectives are those identified by executive management as organizational goals (i.e.: the organizational objective is to improve customer support and services).

Capability Objectives

- DMS is fully mapped to and aligned with the high-level organizational objectives.
- DMS is approved by the executive committee and relevant stakeholders.
- Process is established to ensure the future alignment of the DMS to organizational objectives.

1.1.3. The DMS includes an established mechanism for approval

Capability Objectives

- Mechanism for capturing feedback from relevant stakeholders exists.
- Feedback has been collected and incorporated into the DMS.
- DMS has been reviewed and approved by relevant (named) stakeholders.

1.1.4. The DMS is enforced and audited for compliance

Capability Objectives

- Audit has reviewed and approved the DMS.
- Audit has determined that its implementation can be enforced via existing corporate audit examinations.
- Audit is actively conducting examinations of data programs to ensure adherence to the DMS.

1.2. High Level Business Requirements are Captured, Prioritized, and Integrated into the DMS

1.2.1. High level business requirements have been documented and used to create the DMS

Capability Objectives

- High level business requirements for critical business lines and corporate functions have been documented.
- High level business requirements for critical business lines/corporate functions have been incorporated into the DMS.

1.2.2. **Requirements incorporated into the data management strategy have been prioritized and approved by identified stakeholders.**

Capability Objectives

- Business requirements (incorporated into the data management strategy) have been reviewed, prioritized and approved by identified stakeholders.
- Regular requirements review cycles have been established.

1.3. The DMS Defines the Importance of Identifying, Prioritizing and Assuring the Appropriate Use of Authorized Data Domains.

1.3.1. **The DMS defines the requirement to identify, inventory and prioritize authorized data domains**

Data Domains are logical categories of data that are designated as prioritized factors of input into critical business functions. For example, "trade data" or "regional sales data" could be designated as authorized data domains. Data domains are identified based on the strategic understanding of business requirements and an awareness of the organizational priorities.

Capability Objectives

- DMS defines the importance of data domain designation.
- DMS defines the need to create and govern the data domain inventory.

1.3.2. **The DMS articulates the importance of establishing enterprise policy to enforce appropriate use of authorized data domains.**

Capability Objectives

- DMS defines the need for established policy about the use of authorized data domains.
- DMS defines the need for governance over the use of authorized data domains.

1.4. The DMS is Aligned with and Mapped to Architectural, IT and Operational Capabilities

1.4.1. **Data architecture concepts have been incorporated into the DMS.**

Data Architecture focuses on the design, definition, management and control of data content. This includes giving data business meaning, describing its metadata, and designing and managing taxonomies and ontologies (See Information Architecture)

Capability Objectives

- Data architecture concepts are defined and incorporated into the DMS.
- Data architecture concepts are aligned with stakeholder plans and roadmaps.
- Data architecture concepts are approved by relevant stakeholders.

1.4.2. **Technology concepts have been incorporated into the DMS.**

Technology concepts refer to the strategy, design and implementation of the physical infrastructure (platforms and tools) in support of the DMS. (See Information Architecture)

Capability Objectives

- Technology concepts are incorporated into the DMS.
- Technology concepts are aligned with stakeholder plans and roadmaps.
- Technology concepts are approved by relevant stakeholders.

1.4.3. **Operational concepts have been incorporated into the DMS.**

Operational concepts include such areas as uptime requirements; business continuity planning; retention and archiving guidelines; defensible destruction requirements; privacy standards; etc.

Capability Objectives

- Operational concepts are incorporated into the DMS.
- Operational concepts are aligned with operational goals and objectives.
- Operational concepts have been approved by relevant operations groups.

1.5. The DMS Requires the Creation of an Effective Governance Program.

1.5.1. **The DMS defines the purpose and objectives for establishing data governance**

Capability Objectives

- The DMS specifies the need for the creation of a data governance program.
- The DMS articulates the purpose, objectives and expected outcomes of the data governance program.

1.5.2. **The DMS describes the data governance target state organizational structure**

Capability Objectives

- The target-state governance program organizational structure is defined in the DMS.

1.5.3. **The DMS describes the governance roles and responsibilities**

Roles and responsibilities of the data management organization as well as the roles and responsibilities of the business-line data executives and data stewards are addressed in the DMS.

Capability Objectives

- The DMS identifies the relevant governance stakeholders.
- The DMS describes the roles, responsibilities and relationships of the stakeholders.

1.6. The DMS Defines How the Data Management Program will be Measured and Evaluated.

1.6.1. **The DMS defines how the data management program itself will be measured**

Program categories include areas such as governance, policies, standards implementation, stakeholder buy-in, etc.

Capability Objectives

- The DMS calls for the development of metrics to track program progress.
- Metric plans are socialized with relevant stakeholders.
- Feedback is received and incorporated into the data management strategy.
- Stakeholders review and approve the metric plans and approach.

1.6.2. **The DMS defines the importance of developing outcome metrics to determine the effectiveness of the data management program.**

Outcome metrics areas include improved data quality, reduction in operational fails, improved discovery, access to critical data, etc.)

Capability Objectives

- The DMS calls for the development of outcome metrics.
- Outcome metric plans are socialized with relevant stakeholders
- Feedback is received and incorporated into the DMS
- Stakeholders review and approve the outcome metrics plans and approach

1.6.3. **The DMS defines the approach needed to evaluate the adherence to the DMS implementation.**

The DMS defines how adherence to the strategy will be measured and tracked. This would include items such as appointment of required resources to support the data management program; adoption of standards; compliance with policy; etc.

Capability Objectives

- The DMS identifies the approach to measuring adherence to the Data Management Strategy.
- The approach is socialized with relevant stakeholders.
- Feedback is captured and incorporated into the DMS.
- Stakeholders review and approve.

1.7. The Data Management Strategy Calls for the Creation of a Communication and Training Program

1.7.1. **The DMS describes the importance of data management program awareness through the communication of goals and objectives, scope, priorities, policies and standards.**

Capability Objectives

- DMS identifies the need for a communication strategy.
- DMS defines the core components, scope and reach of the communications and training program.

1.7.2. **The DMS details the need for an education and training program to ensure stakeholder understanding, buy-in and compliance to the data management program**

Capability Objectives

- DMS defines the need for training on the purpose and objectives of the data management program.
- The DMS addresses the approaches and methodologies for a comprehensive data management training program.

2.0   THE DATA MANAGEMENT BUSINESS CASE AND FUNDING MODEL

Definition:
The Data Management Business Case is the justification for creating and funding a data management program. The DM Business Case answers the "why" questions. It addresses the "so what" challenges. It articulates the major data and data related issues facing a firm or business function and describes the expected outcomes and benefits that can be achieved through the implementation of a successful data management program.

Data Management Funding Model provides the rationale for the investment in data management, the mechanism to ensure the allocation of sufficient capital needed for implementation and the methodologies used to measure both the costs and contributions derived from the data management program. Together, the Data Management Business Case and the Data Management Funding Model are critical steps needed to ensure program stakeholder commitment and agreement to the overall objectives of the program.

Purpose:
Data Management is no different than any other established business process. It needs to be justified, funded, measured and evaluated. The Data Management Business Case provides the rationale for the investment in data management. It provides clarity of purpose, enabling agreement and support of program objectives from senior executives as well as program stakeholders. The Data Management Funding Model describes the overall framework used to ensure that the objectives and processes of data management become a sustainable activity within the organization.

Introduction:
The Data Management Business Case articulates the benefits of the data management function, in alignment with the objectives defined, communicated and agreed upon in the Data Management Strategy. It discusses both the defensive benefits of the program (satisfying regulatory requirements, improved risk management, improved data governance and control, improved data quality), as well as the offensive benefits of the program (improved customer service, innovative product development, increased revenues, improved market penetration).

The business case is the cost/benefit realization of the set of activities and deliverables expected from data management and should include both the strategic dimensions (i.e. the establishment of trust and confidence in the end-to-end flow of data) as well as the tactical components (i.e. number of transformation processes, the inventory of data-intensive applications) related to program implementation. The strategic aspects of the business case answers the "why the firm is focusing on data management" questions, helps achieve alignment across the stakeholders and is instrumental in reducing disagreements over ownership of the data management program. The tactical business case is easier to translate into "spreadsheet metrics," is specific to the individual firm and helps management understand the costs, benefits and risks associated with the maturity of the data management program.   In all cases, it is essential to link the business case with realistic strategic and tactical measurement criteria and align both of them with the long term sequence plan for the data management program. This enables the organization to understand the total costs associated with implementation as well as maintenance of the data management program and helps ensure that it is sufficiently funded to meet both near term and long range objectives.

The funding model defines the mechanism used to generate and maintain capital needed for the data management program throughout its lifecycle. It establishes the methodology used for cost allocation

among business lines and can be used to help align stakeholders on funding-related issues. In mature organizations, the funding model reflects the individual requirements of the various components of the organization and is integrated with governance to ensure that appropriate oversight and accountability is applied to data management. Verifiable metrics are essential and the metrics must be aligned with tangible business objectives. A well-structured funding model can help avoid debates over business priorities, mitigate internal competition and facilitate open discussions among relevant stakeholders.

There is no single model for funding data management initiatives. The specific model implemented will depend on the dynamics and operational culture of the individual firm. However, the fundamental components of the funding model should always include areas such as: investment criteria/priorities, budget management, delivered versus expected benefits, allocation methodology and capital needed for ongoing management of the program. It is important to recognize that there will be both fixed and variable components to the funding model for data management.

Strong consideration should be given to allocating initial funding as a fixed corporate expenditure. Data management programs that are funded as a capital investment have a much greater rate of long-term success as compared to 'grass roots' funding. Grass roots funding can become mired by competition among business units, is often aligned with a tactical view of data management and frequently reinforces short-term evaluation cycles. An organization can expect its funding model to evolve along with the maturity of their data management program.

Goals:
→ Create a data management business case based on verified input from stakeholders across lines of business that incorporate both strategic and tactical objectives of the data management program

→ Align the business case to the agreed-upon business drivers and organizational objectives

→ Ensure the business case has been socialized and agree to by program stakeholders to ensure commitment and support of the data management objectives

→ Establish a funding model that supports the agreed upon Data Management Business Case and fits with the culture of the organization in order to ensure buy-in from program stakeholders and commitment to sustainability.

→ Create a mechanism to ensure that the business case remains aligned with business objectives as the organization evolves and matures.

Core Questions:
→ Are the strategic goals of the organization reflected in and aligned with the Data Management Business Case and Funding Model?

→ Is the Funding Model sufficient to support the implementation the data management program.

→ Have the funding requirements been translated into the business case and aligned with the objectives, sequence priorities and implementation roadmap of the data management strategy

→ Does the funding model cover all aspects of data management (e.g.: tangible, intangible, special requests, urgent requirements, unique applications, etc.)

→ Is there a defined process with established criteria for determining and verifying the investment required for data management and is it aligned with the business structure, priorities and governance process organization.

Capabilities and Assessment Criteria:

---

2. DATA MANAGEMENT BUSINESS CASE AND FUNDING MODEL

2.1. The Data Management Business Case is Aligned to Strategic Drivers and Tangible Business Outcomes in Collaboration with Defined Stakeholders.

2.1.1. The Data Management Business Case is mapped to and aligned with drivers, requirements and strategy

Capability Objectives

- The DM Business Case is aligned with the business objectives and strategic priorities of the line of business.
- The DM Business Case is mapped and aligned with organizational priorities and objectives.

2.1.2. High level business outcomes are defined and sequenced.

A primary function of the business case is to define the challenges of the current state and to define the pathway to improvement.

Capability Objectives

- Expected outcomes are defined and sequenced.
- Current-to-Target State is defined and articulated.

2.1.3. DM Business Case is socialized and validated by program stakeholders

Buy-in is predicated on stakeholder validation of the viability of the proposed program

Capability Objectives

- The DM Business Case has been socialized to program stakeholders.
- Target objectives have been reviewed and validated.
- Outcomes, benefits, timelines and target thresholds have been reviewed and approved.

2.2. The Data Management Funding Model has been established, approved and adopted by the organization.

2.2.1. The DM Funding Model is matched to business requirements, implementation timelines and operational capabilities.

Capability Objectives

- DM Funding Model is proposed and socialized with program stakeholders.
- Feedback is being collected and incorporated into the model.
- DM Funding Model has been approved by program stakeholders.
- DM Funding Model is reviewed and enhanced as part of the annual funding process to reflect evolving requirements.

2.2.2. The DM Funding Model is aligned with the business process of the organization

Capability Objectives

- DM Funding addresses current year budget cycle.
- DM Funding is mapped to a multi-year implementation plan.
- Data management funding is integrated as a sustainable corporate function.

2.2.3. Implementation of the DM Funding Model is enforced.

<u>Capability Objectives</u>

- Funding is allocated and approved by the lines of business.
- All budgets are reviewed and approved by the Data Management Organization.
- Data Management Organization is empowered to enforce the line of business data management funding allocation in accordance with DM Program objectives.

2.3. The Funding Model can be measured and evaluated against tangible business objectives

2.3.1. Total expense for the Data Management Program is captured, maintained and analyzed

<u>Capability Objectives</u>

- Total expense is captured, maintained and analyzed at both the line of business and organizational levels.

2.3.2. A standard methodology for calculating ROI is established

<u>Capability Objectives</u>

- Methodology for ROI is established by individual lines of business as well as at the organizational level for aggregate evaluation.

2.3.3. ROI is measured, monitored and used for making Data Management Program decisions

<u>Capability Objectives</u>

- ROI is measured, monitored and used for LOB decision making.
- ROI is aggregated at the organizational level and used to influence data management program priorities.

3.0 DATA MANAGEMENT PROGRAM

Definition:
A Data Management Program is an organizational function dedicated to the management of data as an asset throughout an organization. It illustrates how the management of data quality, definition and content supports strategic, business and operational objectives. It reinforces the necessity of orchestration, active collaboration and alignment among diverse stakeholders in order to instill confidence in data as a trusted factor of input into business and operational processes.

Purpose:
The purpose of a Data Management Program is to embed the concepts of data management into the operational framework of an organization on a sustainable basis. The creation of the data management program elevates the importance of data content management and integrates it as a core component of organizational operations. It establishes data management as a sustainable activity and reinforces the importance of managing "data as meaning" across the organization.

Introduction:
The concept of managing "data as meaning" is not always well understood. For many organizations, data is understood as something to process. It is acquired, normalized, stored, processed and integrated into applications. And while data processing is a critical function, data is also designed to be an accurate representation of real and meaningful things (*i.e. an obligation associated with a business deal, ingredients into a production process, identifying credentials of a customer, role performed in an organizational relationship, etc.*). The establishment of the data management program within organizations is designed to ensure the management of data as meaning and help orchestrate the alignment of data precision with data processing capabilities. These are complementary activities that should be viewed as the "factors of production" for information intensive organizations.

The data management program defines the key components that are needed to ensure trust and confidence in data content and provides guidance for its interaction across the organization. The function of data management is derived from an understanding of business objectives and organizational priorities as well as knowledge of how data flows from initiation through validation through enrichment through transformation and into consuming applications. In many environments, data content is understood as a manufactured product and flows through organization as part of a linked process. It is this linked nature of the process that highlights the collaborative components of data management. The goal is to instill a sense of collective ownership of data quality among all relevant stakeholders.

The data management program should be established as a formal, independent and sustainable part of the organization. The lines of responsibility and accountability need to be established. An inventory should be created to ensure the Office of Data Management has access to the appropriate staff resources and functional capabilities in order to deliver the data needed to support organizational objectives. An effective data management program has the strong support of executive management, appropriate governance authority to ensure to implementation of a control environment for data and a well-structured model of how stakeholders will engage on data-related issues. An effectively designed data management program that is flexible enough to accommodate to changing circumstances will help embed the importance of data content management into the culture of the organization.

Goals:

→ Ensure that the Data Management Program is established, communicated and institutionalized as an independent and sustainable activity

→ Ensure that the Data Management Program is staffed to provide sustainable operation

→ Establish the role, responsibility, accountability and authority of the program stakeholders

→ Establish the stakeholder engagement model to ensure consistency in day-to-day operations, interactions and decision making.

→ Establish the structure and process to ensure that executive management support is institutionalized

→ Define the organizational structure and process to ensure stakeholder's program adherence and adoption as well as conflict escalation and resolution.

Core Questions:

→ Is the data management function aligned with the data management strategy and organizational objectives?

→ Does our organization have the right mixture of skills, resources and capabilities to effectively implement and govern the data management function?

→ Does the Data Management Program have the appropriate support from executive management?

Capabilities and Assessment Criteria:

3. DATA MANAGEMENT PROGRAM

3.1. The Data Management Program is Established and Empowered.

3.1.1. The Data Management Program is established.

Data Management Program is established and communicated to all relevant stakeholders

Capability Objectives

- Data Management Program is sanctioned by executive management.
- The role of the Data Management Program is communicated across the firm through formal organizational channels.

3.1.2. The Data Management Program has the authority to enforce adherence and compliance.

Data Management Program must be formally empowered by senior management and its role communicated to all relevant stakeholders.

Capability Objectives

- Data Management Program is operating collaboratively with program stakeholders.
- Data Management Program has the authority to enforce adherence and compliance through policy and documented procedure.

3.2. The roadmaps for the Data Management Program are developed, socialized and approved.

3.2.1. Program roadmaps are defined, developed and aligned with the Data Management Strategy

Program roadmaps define "target state", and describe the steps required to attain. Roadmap topics include, but are not limited to governance structure; content management strategy; infrastructure design; data architecture; etc.

Capability Objectives

- Program roadmaps are developed.
- Program roadmaps are aligned to all components of the Data Management Strategy.

3.2.2. Program roadmaps are socialized and agreed to by Program Stakeholders

It is essential that roadmaps are shared with relevant stakeholders. Working with stakeholders during the development phases invites collaborative feedback and buy-in.

Capability Objectives

- Data Management Program roadmaps are shared with and aligned to the roadmaps of the program stakeholders (i.e.: architecture; technology; operational roadmaps etc.).
- Stakeholders verify and approve Data Management Program roadmap alignment.

3.2.3. Project plans are developed detailing deliverables, timelines and milestones
Once roadmaps are agreed to and approved, they must be translated into tangible mechanisms for delivery. The Data Management Program Office is responsible for the creation, coordination and management of the data management project plans.

Capability Objectives

- Project plans are developed and aligned to program implementation roadmaps.
- Routine program review procedures are in place to track progress of development plans .

3.3. Stakeholder Engagement Established and Confirmed

3.3.1. Identified stakeholders commit and are held accountable to the Data Management Program deliverables.

Data Management often requires participation and cooperation from staff and resources outside the data management program organizational structure. Those identified as relevant stakeholders must be held accountable for on time and on budget project delivery. To strengthen that commitment, performance in support of the data management program should reflect in stakeholder reviews and/or compensation.

Capability Objectives

- Roadmaps and program milestones have been communicated to the program stakeholders.
- Program stakeholders have reviewed program deliverables.
- Stakeholders are in agreement with (and are being held accountable) to program deliverables.
- Program stakeholders are committed to the program deliverables through job description modification and/or through compensation/bonus.

3.3.2. Funds are allocated and aligned to program roadmaps and workstreams

Sufficient funding dedicated to the data management program must be committed to by business, technology and operations. In a mature data management program, the Data Management Office is granted authority to review and approve committed budgets.

Capability Objectives

- Funding has been allocated and aligned to the program roadmaps and workstreams.
- Funding allocations have been reviewed by the data management PMO.
- Funding challenges have been discussed and reconciled.
- Funding levels have been approved and allocated.

3.3.3. Resource plans are aligned with and verified against program requirements.

Proper resource levels with appropriate skillsets must be secured by relevant stakeholders

Capability Objectives

- Resource planning is complete
- Resource plans have been reviewed by the data management PMO.
- Resource challenges have been discussed and reconciled.
- Resources plans have been approved.

3.4. Communication Program is Designed and Operational

3.4.1. Communication Plans have been created, published and implemented

Capability Objectives

- Communication Plans have been developed and shared with relevant stakeholders.
- Program communications are operational and implemented.

3.4.2. Communications channels have been established.

The full spectrum of communications channels (i.e. websites, access portals, reference libraries, documents, training materials, town hall meetings, etc.) need to be used to ensure that stakeholders understand the goals, objectives and processes associated with the data management program.

Capability Objectives

- Communication channels are established.
- Communications are actively being updated and delivered to relevant stakeholders.

3.4.3. Active engagement with industry regulatory bodies is established to ensure communication of program management and development

Capability Objectives

- Communication to Regulators occurs during routine Regulator meetings.
- Proactive communications are taking place with Regulatory bodies.

3.4.4. Active engagement with external industry and standards bodies are in place

Engagement with industry trade organizations, research consortia and standards bodies ensure that the organization is aware of and aligned with the latest trends associated with data management and new developments related to the data management best practice

Capability Objectives

- Stakeholders are kept abreast of changes and events in the data management industry.
- A formal function is established with dedicated resources to actively participate in data management industry activities and events.

3.5. Data Management Routines are Established, Operational and Measured

3.5.1. Routines for support of the data management program have been established

Routines for steady-state operations of the data management program are taking place. Routines include but are not limited to regular stakeholder meetings, planning sessions, status reporting, etc.

Capability Objectives

- Program routines required for operational support have been identified and scheduled.
- Program routines, meetings and working sessions are taking place

3.5.2. Issue identification, prioritization, escalation and conflict resolution are defined and operational

An established escalation process is necessary to resolve conflicts, reconcile priorities and ensure efficient operations. It demonstrates improved service to the organization, promotes the benefits of an established data management program, and is an important operational routine expected of audit and regulatory reviews

Capability Objectives

- Issue Management routines (meetings; check points; etc.) are defined.
- Issue Management routines are documented and operational.

3.5.3. Metrics (i.e.: KPIs, KRIs) are defined and used to track Program progress.

Capability Objectives

- Program tracking metrics have been designed.
- Program tracking metrics are being captured and reported.
- Program metrics are being analyzed and incorporated into program modifications.

4.0   DATA GOVERNANCE

Definition:
Data Governance is the backbone of a successful Data Management Program. Data Governance is the process of setting standards, defining rules, establishing policy and implementing oversight to ensure adherence to data management best practices. Governance is the formalization and empowerment of the Program to ensure propagation and sustainability throughout the organization.

Purpose:
The purpose of Data Governance is to formalize Data Management as an established business function. Data Governance establishes the rules of engagement, drives funding and prioritization, enforces compliance. Data Governance defines the guidelines for data movement, which defines how data will be acquired, persisted, distributed, appropriately used, archived and/or defensibly destroyed. Data Governance defines oversight by establishing control guidelines, approval processes and evaluation of adherence to policies and procedures. Data Governance ensures that data management principles are fully defined, stakeholders are identified and empowered and adoption is achieved. Governance also ensures that technology, business and operations functions are held responsible and accountable for the maintenance, quality and proper use of data throughout the organization.

Introduction:
Governance is the key to successful data management.   It establishes lines of authority and ensures that the principles of data management can and will be implemented. It establishes the mechanisms for stakeholder collaboration and defines the organizational structure by which the data program will be managed. The governance infrastructure determines where the program resides in the corporate hierarchy, helps manage stakeholder expectations, ensures the adoption of policies and standards, articulates the mechanism for conflict resolution, ensures adequate funding and sets the methodology for measuring data management progress.

Governance over the data management program is multidimensional and includes activities related to strategy, operations, data architecture, IT implementation, data quality and procurement.   It is not created as a steady state activity but will mature and evolve over time. And while the most appropriate structure will vary across organizations, a clear mission with links to tangible business objectives as well as a mechanism for realignment is essential for long term success. For example, domain councils might exist to oversee the intersection of business, technology, and operations. Governing boards might be created to establish business data priorities and resolve conflicts. Tactical groups might exist to manage workflow, perform data reconciliation, address quality of critical data attributes, perform business analysis and provide triage to resolve pressing business challenges with data. All of these components need to be linked into an overall framework if governance is going to successfully embed data management concepts into the culture of the organization as well as manage implementation.

The organizational model for data governance establishes the mechanism by which the data management program is managed, funded and implemented. It defines the management hierarchy and accountability structures for the data program including how people and processes interact. The key objectives are to ensure that the principles of data management are defined and adopted across the organization; that the mechanisms are in place to ensure sustainable funding; and that stakeholders are aligned on the collaborative nature of data management. Executive sponsors are essential to ensure that data governance is successful. Sponsors need to be engaged in both the objectives and structure of the data management program from its inception. The executive mandate helps establish shared expectations and promotes confidence that program objectives are a high priority despite any disruption created to business priorities

and operational structures. Executive sponsors have a critical and active role in managing expectations and in establishing a functional mechanism for addressing competing priorities.

In order to implement governance, the organization needs to ensure that the deployment plan will be effective within their business environment. The governance structure can be used to prevent attempts to "boil the ocean," provide a mechanism to limit selling of obscure technical concepts that don't mean anything to business users, help avoid finger pointing, and minimize environments where stakeholders are put on the defensive. After the initial implementation, the governance framework itself needs to be evaluated, measured and adjusted based on business reality and to ensure that it is fully integrated into operations.

One of the core functions of data governance is to manage the staffing requirements needed to implement the processes and technologies associated with sustainable data management. This should be accompanied by a formal inventory of resource requirements and aligned with the data lifecycle. And since it is not always necessary (or possible) for all essential staff resources to be "owned" by the data management organization, governance is needed to implement the strategy for resource sharing. Resourcing for data management combines IT knowledge, business experience, and data management expertise. The goal is to align skill sets with resources to identify "natural" candidates for new data roles.

The components of data management governance also need to be closely aligned with criteria for measuring the value of data management against defined objectives. Measurement criteria can be used to evaluate the gap between actual and expected value (disparity); the relationship between data management variables (correlation) and the measurement of the data program against objectives (performance). These can be translated into practical measurement criteria such as the cost of correcting mismatches, the time spent on data reconciliation, opportunities for systems consolidation, reduction in the number of transformations, responsiveness to customers, acceleration of business, reduction in operational risk, etc.

Standing Up Your Governance Program

Although this may differ from organization to organization, there are generally 4 steps that are needed in order to establish an effective data management governance program.

1. *Establish the Governance Structure:*
   The aim of the governance structure is to identify and organize the critical stakeholders and link them to the necessary data management support components. In order to implement governance, the organization needs a formal deployment plan to ensure that the governance structure, organizational model, and oversight mechanism will work within the business environment. Interacting with executive management to ensure that adequate funding for data management is in place is critical to ensure that governance is successful.

2. *Implement Policy*
   Formalizing policy is the foundation for Data Governance. Policy addresses how data is gathered, maintained, delivered and utilized. For policy to be effective, it must be enforced and made auditable across the enterprise.

3. *Develop the Operate Model:*
   The operate model must be implemented and deployed to ensure that the data management principles are fully defined, adopted and adhered. The model provides guidance for managing the structure and activities of the data governance program. The model defines the controls, checkpoints and tollgates required, and establishes formal approval processes for the program.

4. *Monitor and Measure:*
   Formal process for adequately monitoring and measuring the effectiveness of the data management program must be deployed to ensure the program is meeting its stated objectives. The program must be evaluated to ensure ongoing consistency with organization policy, and alignment with business strategy. Continuously measuring the program is essential. Metrics-based measurement criteria should be developed and used to track the progress and health of the program. Measurement criteria can include areas such as: measurement of compliance to policy and standards; the cost of correcting mismatches on trade repairs, the time spent on reconciliation, consolidation and better use of existing data sources, reduction in the number of transformations, consolidation of redundant systems, responsiveness to customers, acceleration of business, operational risk, etc.

Goals:

→ Establish executive sponsorship for the program. Communicate purpose and objectives.

→ Establish a functional data management organizational structure with clear role definitions, responsibilities, and accountabilities for data management resources.

→ Establish governance implementation procedures to ensure compliance with policies, processes, standards and resources. Ensure that the structure provides for program oversight, policy enforcement, and issue escalation.

→ Develop comprehensive and achievable policies and procedures.

→ Define clear lines of authority and responsibility for decision-making as well as mechanisms for enforcement of data management based on operational constraints.

→ Ensure that appropriate resources have been allocated to ensure that data governance is effectively implemented.

→ Develop and implement a uniform process for establishing a comprehensive set of metrics. Ensure stakeholder collaboration in the development and use of metrics for meeting data management measurement criteria.

→ Formalize consistent reporting of metrics to identify the progress, health and benefits of the data management program.

Core Questions:

→ Have the data management policies been defined, developed and validated with key stakeholders?

→ Has a governance structure been established? (Stakeholders identified; charters written; responsibilities assigned, etc.)

→ Are there mechanisms in place for issue escalation and resolution?

→ Are there mechanisms in place for establishing and resolving prioritization issues among stakeholders?

→ Are the appropriate executives identified and engaged?

→ Has the methodology to ensure compliance with established policies, processes and standards across the full data lifecycle been defined?

→ Is the funding model and resource strategy sufficient to support the objectives of the data management program?

→ Have the metrics been validated by stakeholder criteria, aligned with business objectives and collected in a timely manner?

→ Are the metrics specific and achievable (actionable) within your organization to improve data management and meet objectives?

Capabilities and Assessment Criteria:

---

4. DATA GOVERNANCE

4.1. Data Governance Structure is Created.

4.1.1. Data Management Office (DMO) is created.

The Data Management Office refers to the centralized organization responsible for championing the data program

Capability Objectives

- DMO is designed and planned**.**
- DMO is approved and charted**.**
- DMO is created.

4.1.2. The DMO has an executive owner.

A senior executive (ex: Chief Data Officer) must be appointed and given full authority to run the DMO

Capability Objectives

- Need for executive owner is recognized, socialized and communicated.
- Executive owner is hired or appointed.
- Duties and authority of the executive owner have been communicated to all relevant stakeholders.

4.1.3. The data governance plan is created

A comprehensive governance plan needs to be built in collaboration with critical stakeholders

Capability Objectives

- Data governance plan is drafted and aligned to operational objectives, priorities and culture.
- Data governance plan has been shared with relevant stakeholders.
- Data governance plan has been reviewed and feedback has been incorporated into the final version.
- Data governance plan is approved.

4.1.4. Program Office (PMO) is established and staffed with required skill sets.

The data program will require the coordination of many projects across a firm or division. Resources may be shared. It is important that a PMO is established and appropriately staffed with adequate resources to manage the required workload.

Capability Objectives

- PMO is approved and chartered.
- PMO is staffed.
- PMO is authorized to ensure and enforce alignment of projects to data management policy and standards.

4.1.5. Enterprise governance structure is designed and implemented

Enterprise governance structure refers to the organizational construct across the enterprise. Individuals must be appointed in business lines and control functions and given the responsibility of data management within those verticals, preferably, reporting into the COO or business leader within that group

Capability Objectives

- Governance structure has been defined, documented and shared with relevant stakeholders.
- Organizational governance structures have been implemented.
- Working committees are established with written and approved charters.
- Stakeholders have been appointed.
- Stakeholder roles and responsibilities have been communicated.
- Stakeholders are held accountable for their participation in the data management program (i.e. via performance reviews and compensation considerations.

4.2. Policy and Standards are Documented, Shared and Approved

4.2.1. Policy and standards are written and complete

Policy and Standards define how business, technology and operations control data including how data is acquired, managed, maintained and delivered throughout an organization. P&S must be developed in partnership with stakeholders to ensure buy-in as well as alignment with existing strategies and controls. Although P&S can vary, most will contain rules and guidelines pertaining to data ownership; data definition, data lineage, metadata, data quality; data access; permissible use; data sourcing and controls.

Capability Objectives

- Policy and standards are developed in collaboration with (business, technology and operations) stakeholders.
- Policy and standards are complete and verified
- Policy and standards are in alignment with Data Management Strategy.

4.2.2. Policy and standards have been reviewed and approved by relevant program stakeholders

Policy and standards must be shared and reviewed by relevant stakeholders to ensure agreement, alignment and buy-in. Policy and standards are critical elements and should be subjected to a rigorous challenge process by stakeholders

Capability Objectives

- Policy and standards have been shared and reviewed by relevant stakeholders.
- Feedback from stakeholders has been incorporated into the final version of the policy and standards.
- Policy and standards have been validated and approved.

4.2.3. Policy and standards have been reviewed and approved by senior executive governing bodies

Policy and standards must be recognized and supported by senior executive management. Data governance must be aligned with (and become a component of) the existing governance structures of the enterprise.

Capability Objectives

- Policy and Standards have been submitted to the organizational governance mechanism for evaluation.
- Policy and Standards have been approved.

4.3. Program Controls are Established

4.3.1. Project "Review and Approval" processes are established

Policy and standards must be enforced in a controlled manner via checkpoints, formal review mechanisms and organizational approval boards. Controlled enforcement must be created to ensure that all new development as well as data access, usage and transmission of data adhere to established policy and standards.

Capability Objectives

- Review and approve processes and responsibilities for data-related projects have been communicated to relevant stakeholders.
- Review and approval processes are operational (includes areas such as "Approval to Build", "Approval to Access", "Approval to Use", "Approval to Send", etc.).
- Review and approval processes are aligned with the control mechanisms of other existing cross-organizational processes (i.e. Change Management policy must referenced, and be harmonized with, data management policy).

4.3.2. Policy and standards are enforceable and auditable

Policy and standards must be supported by established audit processes and routines. Lack of adherence to policy and standards must be elevated as a formal audit issue that must be resolved.

Capability Objectives

- The Data Management Office has the authority to examine and enforce adherence to data management policy and standards.
- Corporate Audit examines and enforces adherence to the data management policy and standards.

4.3.3. Metrics are in place to track program adherence, progress and outcomes

Metrics constitute the empirical evidence required to determine the effectiveness of the data management program. Metrics development is ongoing as new business processes are developed and aligned to the data governance P&S. A successful metrics program will not only capture, aggregate and report metrics, but will also affect program change based on metric evaluations.

Capability Objectives

- Metrics and thresholds are established.
- Metrics are tracked and reported to relevant stakeholders.
- Metrics are tracked and reported to executive management.
- Metrics inform and drive program decisions and modifications.

4.3.4. Formal training programs have been designed and implemented

Behavior and culture change are required for effective data management. Formal training is needed to ensure those with data responsibility are operating in accordance with established P&S.

Capability Objectives

- Training programs are designed and operational.
- Training is mandated as an operational requirement.

4.4. Program Governance is Operational

4.4.1. Meeting routines are established

Governance and control routines (meetings, minutes, actions items, etc.), must be documented and incorporated into practice

Capability Objectives

- Meeting routines have been defined and scheduled.
- Meeting routines are operational.
- Groups are meeting and functioning in accordance with their individual charters.

4.4.2. Requirements are captured and prioritized

Business requirements are an output of the established routines. Sustainable processes are in place to capture, review and verify business requirements.

Capability Objectives

- Routines to capture requirements are established.
- Prioritization process is established and adhered to by relevant stakeholders.
- Prioritization processes for business requirements are designed, rationalized and approved by stakeholders.
- Priorities are reviewed and enhanced on a regular schedule.

4.4.3. Funding Model is Operational

Funding model is implemented across the data management program. The funding model is likely to be multi-year and must be built into the firm's annual funding review cycle. (see business case & funding model process area)

Capability Objectives

- Funding model is operational.
- Funding model is repeatable and aligned to organizational funding cycles.

4.4.4. Escalation Procedures are developed and documented

Formal escalation procedures must be agreed to and documented. Escalation procedures are the mechanism used by the organization to address critical decisions and resolve conflicts.

Capability Objectives

- Escalation procedures have been defined and documented.
- Procedures have been reviewed and approved by relevant executive management and organizational governance bodies.

4.5. Content Governance is Established

4.5.1. Authorized data domains have been identified and inventoried

Authorized data domains are logical representation of a category of data that supports a business function (e.g. "trades" is a data domain that supports capital markets). It is imperative that these strategic data assets are identified and inventoried to ensure their proper use in critical applications

Capability Objectives

- Authorized data domains are identified.
- Authorized data domains are declared.
- Authorized data domains are inventoried.

4.5.2. Critical Data Elements (CDEs) have been identified and inventoried

CDEs refer to the individual data attributes that are used to support critical business functions. CDEs must be identified and catalogued to ensure evidence of proper sourcing, lineage and usage.

Capability Objectives

- CDEs have been identified and inventoried.
- CDE sources have been documented.
- Approved business definitions have been assigned.
- Data lineage has been documented and validated.
- CDEs (and their lineage) are maintained in accordance with data management policy and standards.

4.5.3. Data domain taxonomies have been developed and are actively maintained

Taxonomies define how things relate. Data taxonomies define relationship of elements within a data domain. Taxonomies are critical to establishing common definition and language of data across an enterprise and are required to ensure data's proper use.

Capability Objectives

- Data domain taxonomies are defined.
- Authorized data domains are verified by business subject matter experts.
- Authorized data domain taxonomies are been published and are being used by upstream/downstream systems as data is shared across business processes.
- All new business development use established authorized data domain taxonomies.
- Internal taxonomies are aligned with (and cross referenced to) global standards.

4.5.4. Unique and precise data identification schemes and methodologies have been defined, applied and are in use.

Data Identification schemes and methodologies are used to ensure precise identification of data factors of input. Customer ID; Legal Entity ID; Product ID are examples of unique identification. Establishing ID methodologies are critical for data aggregation, classification and analysis. Unique identification is a foundational concept and is emerging as a required component for regulatory reporting and risk analysis.

Capability Objectives

- Identifiers have been defined for critical business entities (e.g. product; customer; account; etc.).
- Internal entity IDs have been assigned, published and are being used across business processes.
- Internal IDs are aligned (and cross referenced) to industry standard identifiers.

4.5.5. Data classifications are defined and assigned

Data classifications are critical for control and analysis of data. Data classifications are critical to establishing standard treatment of data across an enterprise and for aggregating data for analytical purposes.

Capability Objectives

- Data classifications have been established, assigned to data domains and verified by stakeholders.
- Data classifications are adopted and implemented in systems.
- Data classifications dictate how data is to be handled throughout the business process (e.g.: classification of data denotes privacy treatment, info-security treatment, masking, encryption, risk analysis, etc.).

4.6. Technology Governance is Established.

4.6.1. Platform governance is established

Technology defines and governs how databases and data warehouses are approved, developed and deployed. Technology approach needs to be aligned with the Data Management Strategy.

Capability Objectives

- Procedures for platform governance are defined and developed by IT and are aligned to the Data Management Strategy.
- Platform governance is implemented and operational.

4.6.2. Data distribution governance is established

Technology defines and governs how data is distributed across the network.

Capability Objectives

- A data distribution strategy and governance is defined by IT.
- The data distribution strategy is aligned with the objectives of the Data Management Strategy.
- Data distribution governance is implemented and operational.

4.6.3. Data storage governance is established

Technology defines and governs how data is persisted, archived, restored and defensively destroyed, in alignment with business objectives, the Data Management Strategy, and Legal and Compliance considerations. Storage includes online, archive, cloud and other 3rd party storage medium.

Capability Objectives

- Data storage strategy and governance is developed by IT.
- Data storage strategy and governance is aligned with business, data management and legal and compliance objectives.
- Data storage strategy and governance has been reviewed and approved by relevant stakeholders.
- Data storage governance (people, process, technology) has been implemented and is operational.

4.6.4. BI, ETL and data tool governance is established

Technology defines and governs the technology stack for BI (Business Intelligence), ETL and other data related tools. Data tools include but are not limited to discovery tools, data quality tools, data profiling tools, metadata tools, lineage tools, etc.)

Capability Objectives

- Technology defines the permissible technology stack for ETL and related data tools.
- ETL and data tool governance is implemented and operational across all technology development teams.

4.7. Cross-Organizational Enterprise Data Governance is Established

4.7.1. Data Governance is aligned with Information Security Policy

Capability Objectives

- Data management policy and standards are aligned with Info-Security policy and standards.
- Cross-organizational dependencies are formally recognized and reflected in each groups' policy and standards.

4.7.2. Data governance is aligned with Privacy and Cross-Border Policy

Capability Objectives

- Data management policy and standards are aligned with Privacy and Cross-Border policy and standards.
- Cross-organizational dependencies are formally recognized and reflected in each groups' policy and standards.

4.7.3. Data governance is aligned with external (3rd Party) data usage policy and standards

Firms routinely share their data with 3rd party entities (vendors, service providers, reporting agencies, etc.). Policies and standards are required to govern what data can (and cannot) be shared, what approvals are required to permit external use of data and how data delivered to 3rd parties will be protected (in alignment with corporate information security standards).

Capability Objectives

- Data management policies and procedures for 3rd party data usage have been developed and aligned with business objectives, data management strategy, privacy policies, information security policies, and permissible data usage policies.
- 3rd Party data governance policies and standards are implemented and operational.
- Cross-organizational dependencies are formally recognized and are reflected in relevant groups' policies and standards.

4.7.4. Data governance is aligned with Legal and Compliance Data Policy

Data Management strategy and governance must be aligned with legal and compliance data policies not already discussed.

Capability Objectives

- Data Management policies and standards are aligned with Legal and Compliance data policy and standards.
- Cross-organizational dependencies are formally recognized and reflected in relevant groups' policy and standards.

INFORMATION ARCHITECTURE

There are three architectural disciplines that are important for firms to understand in order to operate cohesively and achieve their organizational data management objectives. Business Architecture, which define business processes; Data Architecture, which defines the meaning and relationships of data within those business processes and Technology Architecture, which defines how technology infrastructure (platforms, tools and applications) work together to enable access and use of data.

Information Architecture is the result of combining the disciplines of Data Architecture and Technology Architecture to support the goals and objectives of the business. There are four elements that need to be understood in the context of Information Architecture.

- Semantics: Semantics refers to the adoption of precise, shared and consistent meaning. At the heart of the Data Management challenge has been the inability to harmonize disparate data across an enterprise or a business function because the meaning of data has not been standardized. Semantics is the discipline of assigning unambiguous meaning to data throughout the data lifecycle.

- Conceptual Data Model: Conceptual data models describe data at its highest level, identifying the critical data objects needed to satisfy a business objective as well as defining their relationships to one another.

- Logical Data Model: The Logical Data models is a fully-attributed conceptual model that has been abstracted from any physical implementation. The logical model represents the business requirements in terms of what is needed to satisfy the objectives of the business function.

- Physical Model: The physical model is the instantiation of the meaning, relationships and attributes of data into a physical implementation.

The following section will look at Information Architecture through the lens of its component parts – Data Architecture with its focus on content meaning and Technology Architecture with its focus on platform, tools and applications.

5.0   DATA ARCHITECTURE

Definition:
Data Architecture speaks to the design, definition, management and control of information "content". Data Architecture identifies data domains, documents metadata, defines critical data elements, establishes taxonomies and models ontologies that are critical to ensuring that the meaning of data is precise and unambiguous and that the usage of data is consistent and transparent.

Purpose:
A Data Architecture function establishes consistency in definition and use of data throughout an organization. Adhering to a prescribed data architecture forces business and technology to take the necessary steps to define and document data meaning, define the appropriate use of the data, and to ensure that proper governance is in place to consistently manage "data as meaning" on a sustainable basis.

Introduction:
Data exists throughout an organization across all facets of business operations. The design of a firm's Data Architecture is based on a comprehensive understanding of business requirements. Unraveling the business process informs how data should be identified, defined, modeled and related. Technology Architecture then dictates how the data architecture design is instantiated into physical repositories in order to provide optimized access, security, efficient storage management and speed of processing.

In order to establish a successful Data Architecture program, there are a number of specific architectural 'steps' that must be developed and adhered to. First is to understand the scope of data needed to satisfy the business requirements. The scope of data generally falls into two categories: (1) identification of logical domains and (2) identification of the physical repositories.

- Identification of Logical Domains of Data
  Logical domains of data represent the data (not the databases) that are needed to satisfy the business requirements. Logical data domains are grouped into three categories:

  1. Reference Data Domains ("nouns"): These describe the formal or contractual attributes of a business object (i.e.: products, instruments; customers, legal entities; counterparties; etc.)
  2. Transaction Data Domains ("verbs"): these describe the actions associated with an event (sale; trade; deal; payment; etc.)
  3. Derived Data Domains ("adjective"): these describe the newly created, quantitative values that result from an aggregation or analytical operation of reference and/or transactional attributes (i.e.: calculated balances; exposure metrics; demographic calculations; etc.)

- Identification of Physical Repositories
  Underlying the logical data domains are multitudes of physical (often overlapping) repositories of data that will map into the logical data domains. Identification of these underlying physical repositories is a critical step towards minimizing the complexity of legacy environments, reducing replication, better understanding data lineage, assigning data ownership and assessing data quality.

Once the domains (and their underlying physical sources of data) have been identified, precise business definitions (common semantic language) for the identified data entities must be assigned and agreed upon

by critical stakeholders. Data Architecture is about managing meaning. The importance of assigning precise definitions in the context of business reality (relationships), the creation of a shared 'data dictionary' and getting the buy-in from both upstream and downstream users cannot be minimized. Without this common understanding of data attributes (aligned to business meaning), Data Architecture will struggle to succeed, the risk of inappropriate use of data will increases and the ability to 'share' data across an enterprise with confidence will be hindered.

The next step in addressing data architecture is to define *data taxonomies and business ontologies*. Data taxonomies define how data entities are structurally aligned and related. For each officially designated data domain that is identified, inventoried and deemed critical, a taxonomy must be defined, maintained and mandated for all systems using this data as input into their business functions. With critical business function taxonomies defined and in place, the organization needs to model the relationships between taxonomies into a business *ontology*. Ontologies represent the relationships and knowledge of multiple related taxonomies across functional domains.

Semantics, taxonomies and ontologies define and relate the content of data in order to enable the organization to realize its maximum value in a consistent and controlled manner. Once the content is defined, it needs to be precisely described as *metadata*. Metadata falls into three categories: descriptive metadata, structured metadata and administrative metadata.

- Descriptive metadata describes attributes used for discovery and identification (i.e.: author; title; source).
- Structural metadata describes how attributes are created or derived (ex: a derived attribute would describe what attributes that were used to derive its value).
- Administrative metadata provides information related to the creation, classification and/or appropriate use of data. Administrative metadata would include information like "NPI"-Non Public Information; data access entitlements; archive and retention requirements, etc.

Goals:

→ Data Architecture defines common meaning of data

→ Common meaning is driven by business stakeholders

→ Relationships between and among data attributes is based on business requirements

→ Data is designed logically, abstracted from physical implementation

→ Data Architecture informs physical implementation

Core Questions:

→ Are business stakeholders driving content definition?

→ Are policies in place to govern the creation and maintenance of data attributes and relationships?

→ Are governance procedures in place to ensure adherence to established data architecture standards?

→ Are design reviews in place and required to ensure enhancements and new development are utilizing standard data architecture definitions?

→ Is adherence to data architecture standards auditable?

Capabilities and Assessment Criteria:

---

5. DATA ARCHITECTURE

5.1. Identify the data (logically and physically)

5.1.1. Logical domains of data have been identified, documented and inventoried.

*Logical domains of data represent the data (not the "databases") that are needed to satisfy the business requirements. Logical data domains fall into three categories - reference data; transactional data; and derived data. Identification of these domains must be driven by the Business from the perspective of "what **data** is needed to perform the required business functions?"*

Capability Objectives

- Business stakeholders have been selected to drive the identification of the logical data domains.
- Logical data domains have been identified and prioritized.

5.1.2. Underlying physical repositories of data have been identified, documented and inventoried

Underlying the logical data domains are physical (often legacy) repositories of data that will feed the logical domains.

Capability Objectives

- Underlying physical repositories linked to the logical data domains have been identified
- Identified repositories have been inventoried and the inventory is actively maintained.

5.2. Define the Data (semantically and structurally)

5.2.1. Attribute level "business" definitions are defined, documented and approved by relevant stakeholders

Business definitions are non-technical descriptions of data attributes that are based in contractual, legal and/or business facts

Capability Objectives

- Business definitions are documented, verified by users and assigned to physical repositories.
- Business definitions are assigned to fully-attributed conceptual models.
- Authorized data domains are fully populated with business definitions.
- The conceptual mapping is agreed to by relevant stakeholders.

5.2.2. Taxonomies and ontologies are created, documented, maintained and governed

Data taxonomies define how data entities are structurally aligned and related. For each officially designated data domain that is identified, inventoried and deemed critical, a taxonomy must be defined, maintained and mandated for all systems using this data as input into their business functions

Capability Objectives

- Taxonomies are defined for fully-attributed conceptual models and agreed to by relevant stakeholders.
- Taxonomy relations are captured and documented into domain ontologies.

5.2.3. Metadata is defined

Metadata must be captured and inventoried into a metadata repository so it is usable by all relevant development teams

Capability Objectives

- Metadata from the physical repositories is captured and inventoried.
- Metadata for fully-attributed conceptual models is captured and inventoried into a metadata repository, and is being used by relevant stakeholders
- Metadata is rationalized across taxonomies and ontologies

5.3. Govern the Data (establish sustainable data architecture governance)

5.3.1. Data architecture governance procedures are in place and aligned with business governance processes

Alignment to Business processes include: Business Process Definition; Operations procedures; 3rd party contract specifications; etc.

Capability Objectives

- Data governance is aligned with business processes to ensure semantic definitions, taxonomies and CDEs are properly assigned, maintained.

5.3.2. Data architecture governance procedures are in place and aligned with technology

Alignment to Technology processes include: design reviews, approvals to build approvals, validation of appropriate usage approvals, permit to deliver approvals, etc.

Capability Objectives

- All technology development is required by governance policy to follow data architecture standards
- All technology development use established data architecture elements.

6.0   TECHNOLOGY ARCHITECTURE

Definition:
Technology Architecture refers to the strategy, design and implementation of the physical architecture in support of the defined data architecture. Technology architecture defines the platforms and the tools and how they need to be designed for maximum efficiency in support of the data management strategy. The purpose of technology architecture is to define how data is physically acquired, moved, persisted and distributed in a streamlined and efficient manner. Physical data proximity, bandwidth, processing time, backup and recovery, archiving, etc. are all important elements of a mature technology architecture.

Purpose:
The efficient and effective movement of data is critical to business operations. Technology architecture determines how data, tools and platforms operate in collaboration to satisfy business requirements. The proper alignment of these components dictates application efficiency and system processing speed. This enables firms to control costs and achieve infrastructure scalability and elasticity which are the characteristics of an enterprise infrastructure that is designed for long-term implementation success. Technology Architecture is articulated in the technology architecture roadmap. The technology architecture roadmap defines the target state infrastructure and provides guidelines for implementation. The roadmap further defines the technology governance and controls that are needed to ensure compliance across the enterprise.

Introduction:
Information Architecture is the combination of both data architecture (content) with technology architecture (implementation). Data architecture should not dictate technology. Technology is the responsibility of the technology department. However, Data Architecture does *inform* technology. Data Architecture captures the information requirements of the business and translates them into the "what, where and when" of data – what data is needed; where is it to be delivered and by when. Technology Architecture is the *enabler* and defines the plan and roadmap for implementation.

There are four areas of technology architecture that are critical to a successful data management program.
1. *Database Platforms*: Technology Architecture defines acceptable data platforms for enterprise use. Enterprise-class database platforms, appliance technologies, distributed computing, and in-memory solutions all need to be defined, communicated and governed by technology architecture.
2. *Tools*: Often one of the biggest expenses and source of inconsistent handling of data is the proliferation of multiple, disparate data management technology tools within an organization. Technology Architecture must define the allowable tool stacks – what BI (Business Intelligence) tools, ETL (extract, transform, load) tools and various discovery tools are permitted for use within the organization.
3. *Storage Strategy*: Technology architecture must define how firms will store and maintain its data. A component of the target-state storage strategy is the determination of how data and data costs will be maintained, how and what data will be stored (including decisions about the use of internal versus external cloud technology), how data will be archived and retained, and how data will be defensibly destroyed/removed from the firm's infrastructure.
4. *Operational Risk Planning*: A sound technology architecture addresses operational risk, business continuity and disaster recovery strategies. Data is the 'life-blood' of a firm and needs proper planning to ensure that data flows to all parts of a firm even in the face of events that interrupt business continuity.

And finally, all of the above elements of a sound technology architecture must be supported by a strong technology governance operate model. Policies must be in place, agreed to by all technology and business stakeholders, supported by executive management, and subject to internal audit scrutiny and adherence. Without governance, technology will grow and develop uncontrolled and lead to inefficiencies and security issues putting data quality at risk.

Goals:

→ Data Architecture defines target-state infrastructure in support of data management

→ Tool selection and implementation is simplified, reducing complexity and cost

→ Storage strategy is developed consistent with the objectives of business while controlling cost and risk

→ Operational risk architecture is implemented to ensure continuous flow of data to critical business functions in the event of an outage incident

Core Questions:

→ Is technology architecture being driven by business requirements?

→ Are policies in place to govern the selection and use of technologies throughout the organization?

→ Are governance procedures in place to ensure adherence?

→ Are design reviews in place and required to ensure enhancements and new development are utilizing standard technology architecture definitions?

→ Is adherence to data architecture standards auditable?

Capabilities and Assessment Criteria:

---

6. TECHNOLOGY ARCHITECTURE

6.1. Data Platform Strategy Defined and Governed

6.1.1. Technology architecture strategy is defined and agreed to by relevant stakeholders.

Capability Objectives

- An integrated technology architecture strategy is designed, socialized and agreed by relevant technology, business and senior executive stakeholders.
- An integrated architecture strategy is supported and enforced by corporate audit policy.

6.1.2. An actionable roadmap is developed and adopted for implementation of the technology architecture

For a technology roadmap to be sustainable, it must have a budget commitment over the life of the designed roadmap.

Capability Objectives

- A multi-year technology architecture roadmap has been developed.
- The roadmap adheres to the approved technology architecture strategy.
- Budgets have been developed and approved as well as built into the firm's budget cycle processes.

6.1.3. Platform governance structure and processes are in place.

Capability Objectives

- Integrated governance structure and policies are in place, operational and in alignment with the data management strategy.
- All enhancements and new development are subject to architectural platform design review and approval.

6.2. Data Technology Tool Stack Defined and Governed

6.2.1. Technology tool selection strategy is defined and verified by relevant stakeholders

Capability Objectives

- Integrated technology tool strategy has been designed, socialized and agreed to by relevant technology, business and senior executive stakeholders.
- The tool strategy is supported by corporate policy and enforced by Corporate Audit.

6.2.2. Technology tool roadmap is developed and implemented

Capability Objectives

- Integrated technology tool roadmap has been developed in adherence to the technology tool strategy (including guidelines for new development as well as decommission plans for non-standardized legacy tool implementations).
- Budgets have been developed and approved and have been built into a firm's budget cycle processes.

6.2.3. Tool selection governance structure and process is in place and operational

Capability Objectives

- Integrated tool governance structure (with associated policies) are in place, operational and in alignment with the data management strategy.
- All enhancements and new development are subject to tool selection review and approval.

6.3. Data Storage Management Strategy Defined and Governed

6.3.1. Data storage management strategy is defined and agreed to by relevant stakeholders

Capability Objectives

- Integrated storage management strategy has been designed, socialized and agreed to by relevant technology, business and senior executive stakeholders.
- The storage management strategy is backed by Corporate Audit.

6.3.2. Data storage management roadmap is developed and implemented

Capability Objectives

- Integrated data storage roadmap has been developed.
- Budgets have been developed, approved and integrated into the firm's budget processes.

6.3.3. Storage governance structure and processes are in place and operational

Capability Objectives

- Integrated storage management governance structure and policies are in place and operational.
- All enhancements and new development are subject to a review and approval consistent with the defined storage management strategy and roadmap.

6.4. IT Operational Risk Planning in Place

6.4.1. Data infrastructure contingency planning is defined and in place

Capability Objectives

- Integrated IT operational risk management strategy has been developed, socialized and agreed to by relevant technology, business and senior executive stakeholders.
- Integrated IT operational risk management strategy is backed by Corporate Audit.

6.4.2. Operational Risk Governance Structure and processes are in place and operational

Capability Objectives

- Integrated operational risk governance structure and policies are in place and in operation.
- All enhancements and new development are subject to a review and approval of their operational risk plans.
- Operational risk planning is subject to Corporate Audit.

7.0 DATA QUALITY PROGRAM

Definition:
Data Quality describes the degree in which data is fit for purpose for a given business process or operation. Terms such as accuracy, completeness and timeliness are all components of data quality. Data Quality is not a process itself, but is the net result of a chain of processes across the full data supply chain to ensure that data delivered meets the needs of its intended consumers. Data Quality requires an understanding of how data is sourced, defined, transformed, delivered and lastly, consumed.

Purpose:
The Data Quality Program defines the goals, approaches and plans of action to ensure that data content is of sufficient quality to support defined business and strategic goals of the organization. The Data Quality Program should be developed in alignment with business objectives, measured against defined data quality dimensions and based on an analysis of the current state of data quality.

Introduction:
Data quality is a broad conceptual term that needs to be understood in the context of how data is intended to be used. Perfect data is not always a viable objective. The quality of the data needs to be defined in terms that are relevant to the data consumers to ensure that it is fit for its intended purpose. The overall goal of data management is to ensure that users have confidence that the data they are using for decision making accurately reflects the facts the data is designed to represent - without the need for reconciliation or manual transformation.

The organization needs to develop a data quality strategy and establish the overall plans for managing the integrity and relevance of data. One of the essential objectives is to create a shared culture of data quality stemming from executive management and integrated throughout the operations of the organization. In order to achieve this cultural shift, the organization must agree on both requirements and the measurement of data quality that can be applied across multiple business units and applications. This will enable business sponsors, data consumers, and IT to link data quality management processes with objectives such as better risk management, enhanced analytics, better client service and improved operational efficiencies.

Data quality can be segmented into a number of core dimensions including: completeness (*the availability of required data attributes*), coverage (*the availability of required data records*), conformity (*alignment of content with required standards*), consistency (*how well the data complies with required formats/definitions*), accuracy (*the relationship of the content with original intent*), duplication (*the redundancy of records and or attributes*), and timeliness (*the currency of content representation as well as whether the data is available/can be used when needed*). The identification and prioritization of the dimensions of data quality fosters effective communication about data quality expectations and are an essential pre-requisite of the data management program.

Creating a profile of the current state of data quality is an important initial component of the overall data quality initiative and should be performed periodically (i.e. whenever data is transformed). The goal is to assess patterns in the data as well as to identify anomalies and commonalities as a baseline of what is currently stored in databases and how actual values may differ from expected values.   Once the data profile is established, the organization needs to evaluate the current state against data quality requirements (i.e. tolerances and thresholds) as well as against business requirements to determine whether the data is fit-for-purpose.

The underlying purpose of this assessment process is to measure the quality of the most important business attributes and to determine what content is in need of remediation. A robust Data Quality Program (in partnership with business) identifies and declares what data is most important. It is this identification of *critical data elements*, or CDEs that helps inform business and technology in terms of which attributes are prioritized for key business functions. The designation of CDEs helps ensure that the highest level of accuracy and data quality treatment is applied. This is the data that needs to be "cleansed" to meet data consumer requirements. Data cleansing should be performed against a pre-defined set of business rules to identify anomalies that can be linked to operational processes.

Data cleansing should be performed as close to the point of capture as possible. There should be a clearly defined strategy (with owners) for data cleansing to ensure that cleansing rules are known and to avoid duplicate cleansing processes at multiple points in the information management cycle. The overall goal is to clean data once at the point of data capture based on verifiable documentation and business rules as well as to reconcile the processes that allow defective data content into the system. Data corrections must be communicated to (and aligned with) all downstream repositories and upstream systems. It is important to have a consistent and documented process for issue escalation and change verification for both internal originators and data vendors. It is also important to ensure that data meets quality standards throughout the lifecycle so that it can be integrated into operational data stores. This component of the data quality process is about the identification of content that is missing, determination of data that needs to be enriched and the validation of data against internal standards to prevent data errors before data is propagated into production environments.

In order for Data Quality to be sustained, a strong governance structure must be in place to support the data quality activities, ensure compliance to data quality processes and ensure the highest level of organizational support (senior executive management). Data quality processes need to be documented, operationalized and routinely validated via data management reviews and formal audit processes.

Data quality cannot be achieved centrally or monolithically. Enterprise Data Quality requires the commitment and participation of a broad set of stakeholders. Since data quality is the result of a chain of business processes, stakeholders along that chain must be in place, authorized and held responsible for the quality of data as it flows through their respective areas. Data Quality requires coordinated organizational support. Data quality processes and objectives must be part of the operational culture of a firm for it to be sustained and successful.

Goals:

→ Data quality strategy is aligned with business plans and target operating models.

→ Standard dimensions associated with data quality are defined and prioritized by stakeholders.

→ Data quality processes (profiling, assessment, cleansing, and integration) are established and used for all data initiatives across the full systems lifecycle.

→ Data profiling methodologies are standardized, documented and implemented across all critical data stores and repositories.

→ Data quality metrics (tolerances, logic checks, thresholds, duplications, null sets, padding/meaningless spaces, all/no capitals, string length) are defined, documented, aligned with

business requirements.

→ The root causes of data errors are researched. Data quality cleansing and remediation is prioritized based on organizational requirements and business criticality.

→ There is a bi-directional communication mechanism in place with suppliers to improve overall data quality. Data quality processes between data vendors and the organization are documented via SLA's and synchronized.

Core Questions:

→ Is it understood that Data Quality is not an objective unto itself, but an indication of an inefficient business process or broken technology?

→ Is it understood that Data Quality is an cultural shift? Improved data quality touches all aspects of business and technology processes.

→ For a Data Quality Program to be sustainable, training is required. Are the necessary resources (dollars and people) earmarked to implement and operate an Enterprise Data Quality Program?

Capabilities and Assessment Criteria:

---

7. DATA QUALITY

7.1. Data Quality Program is Established.

7.1.1. The data quality strategy and approach is defined and socialized

Data Quality strategy and approach encompasses the "what/how/who" of data quality. It needs to address the scope of the data to be scrutinized and reviewed; how the DQ assessments will be performed (metrics defined) and who will be responsible (defined roles and responsibilities). Data Quality involves cultural change. It is critical that a documented DQ strategy and approach is socialized with relevant stakeholders (technology, business and operations), to ensure awareness, support and commitment.

Capability Objectives

- DQ strategy and approach has been designed and developed
- DQ strategy and approach has been communicated to relevant stakeholders and
- Feedback from stakeholders has been incorporated into the final version of the DQ strategy.
- Stakeholders and Senior Management endorse and support the DQ program and strategy.

7.1.2. Accountable parties have been identified and roles and responsibilities have been assigned.

A Data Quality program requires a network of data stewards and subject matter experts to ensure data is properly captured, processed and delivered. Accountable parties must be identified and their roles and responsibilities must be clearly and unambiguously communicated.

Capability Objectives

- Accountable parties have been identified.
- Accountable parties have been aligned to the organizational data management governance structure.
- Data quality responsibilities have been assigned.
- Individuals are held accountable for the performance of their data quality function via annual reviews and compensation considerations.

7.1.3. The Data Quality roles and responsibilities have been communicated.

The organization needs to be aware of the assigned roles, responsibilities, and authorities for the DQ program.

Capability Objectives

- DQ Governance roles and responsibilities have been defined.
- DQ Governance roles and responsibilities have been aligned with overall governance strategy.
- Roles and responsibilities have been communicated to relevant stakeholders - feedback has been incorporated into the final role definitions.
- Stakeholders and Senior Management endorse and support the defined roles and responsibilities.

7.2. The Current State of Data Quality is Assessed and Remediation Plans are Developed

7.2.1. All relevant data have been identified and prioritized.

In performing a current state data quality analysis, it is important to include all relevant data in order to determine the true state of data quality. Data must be profiled, analyzed and graded. Data remediation plans need to be developed and prioritized. Data Architecture defines the domains of data as "reference data, transactional data, and derived data". Across these domains, two states exist – current and historical. All aspects and all states of data must be considered when performing a current state analysis. Across this scope of data, data can be prioritized based on their relevancy to critical business function. Data elements that are important to prioritized business functions are identified and designated as Critical Data Elements (CDEs). Designated CDSs receive the highest levels of monitoring to ensure the quality of these attributes is maintained. CDE designation is a controlled process. Changes (additions or deletions) to the list of CDEs must be reviewed and approved by business stakeholders

Capability Objectives

- The scope of data subject to the data quality program has been identified (current and historical)
- The scope of data has been prioritized in alignment with the data management strategy and business priorities.
- CDEs are defined, verified, designated and actively maintained.

7.2.2. Data is profiled, analyzed and graded

The scope of data under consideration must be profiled to determine the full spectrum of data quality dimensions (e.g. completeness, timeliness, coverage, conformity, referential integrity, consistency and levels of duplication and redundancy). This analysis must include both a row-based analysis (accuracy of the record) and a column-based analysis (statistical columnar analysis). Metadata must also be reviewed to ensure the description and intended use of data is properly defined.

Capability Objectives

- In-scope data has been profiled and statistically analyzed.
- Metadata has been reviewed and gap analysis has been performed.
- In-scope data has been graded and catalogued.

7.2.3. Data remediation has been planned, prioritized and actioned.

Based on the current state analysis performed, remediation plans must be developed to address the most pressing data quality issues, as well as timelines established for ongoing DQ evaluation and maintenance

Capability Objectives

- Data remediation plans are developed and prioritized
- High priority data remediation is being actioned.
- Timelines have been established for ongoing remediation

7.3. Data Quality Program is Operational

7.3.1. Data Quality 'control points' are in place along the full spectrum of the data supply chain.

A Data Quality Program is developed to quantitatively measure the quality of data as it flows across business and technology processes. Data Quality is governed by setting goals and objectives, establishing control points, determining root-cause, remediating data gaps, and holding the business, data and technology teams accountable for achieving and sustaining the highest data quality standards possible. Data quality control points validate data along the supply chain. Data Quality Controls include the implementation of business rules, establishing workflows, setting data quality tolerances, defining exception handling processes and defining escalation procedures as data moves from data provider to data consumer.

Capability Objectives

- Data Quality control points are in place and fully operational along the data supply chain
- Control remediation procedures are documented and evidenced.

7.3.2. Data Quality Metrics are captured, reported and used to drive data remediation.

Control points capture data quality metrics and produce routine reports for executive management. Metrics are used to track data quality progress and stability and drive data remediation efforts.

Capability Objectives

- Data Quality metrics are captured on a routine basis.
- Data Quality metrics are being reported to executive, business and technology management
- Metrics are being used to drive and prioritize remediation efforts.

7.3.3. Root-Cause analysis is performed

Data remediation is not only about correcting existing data, it is also about determining the root-cause of the data quality deterioration at the source to avoid damaging data in the future.

Capability Objectives

- Root-cause is determined
- Corrective measures to business and/or technology processes are identified and implemented.

7.3.4. Data Quality processes are audited

Data Quality auditing occurs on 3 levels: Quality Assurance (QA) Assessment: Business performs self-assessments based on defined data quality processes and objectives. Quality Control (QC): The Data Management Function preforms a facilitated audit of a business-line's data quality processes and is empowered to enforce the business lines to remediate any gaps found to ensure adhere to data quality best practices. Corporate Audit: Business line data quality processes are subject to corporate audits. Failure to satisfy this review may result in formal escalated audits written against a business line or function.

Capability Objectives

- Data Stewards have performed self-assessment of the business-line data quality processes (QA).
- The Data Management Organization has performed facilitated assessments of business-line data quality operations (QC).
- The Data Management Organization is empowered to force operational teams to remediate gaps found in their operational data quality processes.
- Corporate Audit performs routine examinations of business-line Data Quality procedures.
- Formal Audit Issues are generated if operational gaps be uncovered.

8.0   DATA OPERATIONS

Definition:
Data Operations is the organizational process by which the data assets of a firm are managed in order to realize their maximum value. There are three elements of Data Operations:

1. The orchestration of data management capabilities within a controlled operational model.
2. The stewardship of the Data Management Lifecycle – from source to consumption to disposition
3. The integration of data management into the "information eco-system" (how data management coordinates with other control functions within an organization).

Purpose:
The purpose of Data Operations is to coordinate the people, process and technology of data management into a cohesive operational model. Data Operations defines the mechanisms used to capture requirements, unravel data flows and linked processes and determine how data is to be delivered to the end-consumer. Data Operations supports the Data Management Lifecycle. It ensures that proper resources and controls are in place as data moves throughout its lifecycle journey. And Data Operations ensures collaboration and alignment to cross-organizational control functions. Areas such as Information Security, Data Privacy and Change Management must operate in sync with Data Management in order to ensure data is properly managed across all business functions.

Introduction:
One of the first functions within Data Operations is the orchestration of the data management capabilities and component disciplines. These disciplines have to be aligned to effectively manage data across an organization. Data Operations is the process of aligning all of the capabilities discussed in this model into a consistent operational flow. Each capability has to be properly resourced and prioritized as well as supported by business, technology and senior management.

The successful coordination of these elements is a determining factor in the success of the data management program. It is the responsibility of the data management organization and the Chief Data Officer to structure and coordinate the operational model in order to properly define data meaning, ensure data quality, and deliver data in a timely and efficiently manner. And this must all be evidenced through demonstration of organizational structures, charters, policies and senior management directives.

Data is a core factor of input into business functions and operational processes. The Data Management Lifecycle tracks the progress of data from source … to storage … to maintenance … through distribution … to consumption … to reuse … to archiving … and finally to defensible destruction. The mechanisms used to identify, align and validate the data as factors of input into business functions are derived by reverse engineering existing processes into their individual data attributes and by unraveling the "data assembly" processes used to create the required data sets.

This reverse engineering (or data requirements definition) process needs to be managed with precision to avoid confusion and miscommunication between what the business users truly need for their intended application and what IT professionals need for technical implementation. Data requirements should be modeled, aligned with business meaning, prioritized in terms of how critical it is to the application, verified by all relevant stakeholders and re-verified by stakeholders to ensure that essential concepts are not "lost in translation." This is particularly critical for data that is shared among multiple users and for core data attributes that are used as a baseline for onward expression in operational calculations or business formulas.

For complex applications and for all aggregation-related processes, it is essential to understand and document how the data moves from system-to-system; how the data is transformed or mapped; and how the data is aligned to business definition/standard meaning. Gaining agreement on this "lineage" process is fundamental for ensuring that the results of decentralized or linked processing can be trusted to be consistent and comparable.

The final element of Data Operations is the integration of data management into the "Information Ecosystem" of an organization. The Information Ecosystem is a concept that describes how data is managed collaboratively across all enterprise control functions. Control functions such as Information Security, Storage Management, Legal and Compliance, Privacy, and Vendor Management all have responsibilities on how data is managed. It is imperative that the policies of data management are integrated and aligned with the policies of the cross-organizational control functions to ensure data is being managed consistently and holistically across the organization.

Additionally, Data Operations must align to technology development policies and best practices. Data Management capabilities such as Architecture, Governance and Data Quality should all be integrated into a firm's SDLC (Software Development Lifecycle) processes to ensure that data management considerations are being adequately addressed at the appropriate stages of the development cycle. Nothing should operate in a silo. Operating within an ecosystem recognizes interdependencies and ensures collaboration.

Goals:

→ Data Operations are functional and aligned with Data Management capabilities
→ Data Operations are accountable to the data management organization (or Chief Data Officer) to ensure consistency in how data is managed along the data supply chain.
→ Data Operations follow defined data management capability best practices and are routinely subject to review and audit
→ Data Operations supports all facets of the Data Management Lifecycle
→ The Information Ecosystem is properly governed and controlled. All data entering the system is subject to the collective policies across all control functions.

Core Questions:

→ Are Data Operations units functioning independently within business silos?

→ Does the Data Management Function have accountability for how Data Operations are run?

→ Do Data Operations units support the Data Management Lifecycle?

→ Are Data Operations units subject to routine audits? Have operational gaps been identified and are remediation plans in progress?

→ Does Data Operations collaborate with cross-organizational control function teams?

→ Is data managed differently across control functions?

Capabilities and Assessment Criteria:

---

8.   DATA OPERATIONS PROGRAM

8.1. Data Operations are aligned with Enterprise Data Management Capabilities and Strategy.

8.1.1. Data Operations are aligned to the Data Management Capabilities and Strategy

Data Operations can operate centrally, or federated across an organization. In either model, Data Operations must be aligned with the enterprise data management organization, capabilities and strategy Data Operations must function consistent with the overall data management capabilities and strategy. Data Operations must adhere to operational standards and are subject to routine audit

Capability Objectives

- Data Operations functions in alignment with enterprise data management capabilities and strategy.
- Data Operations is subject to enterprise data management policies and standards.
- Data Operations are subject to Corporate Audit to ensure alignment to policy and standards

8.1.2. Data Operations are accountable to the Data Management Organization

It is essential that Data Operations functions across the Enterprise are aligned and accountable to the enterprise data management function (Data Management Office) to ensure consistency in approach, coordinated prioritization and assurance of proper resource allocation and funding.

Capability Objectives

- Data Operations teams across the enterprise are accountable to the enterprise Data Management Office.

8.1.3. Data Operations follow industry best practices

Regardless of whether Data Operations is centralized or federated, the process and methodology for capturing requirements and designing data solutions must be consistent with and aligned to internal and industry best practices.

Capability Objectives

- Data Operations teams capture data requirements and design data solutions that are consistent across businesses and in alignment with enterprise data management best practices.
- Data Operations teams capture data requirements and design data solutions that are consistent across businesses and in alignment with industry best practices

8.2. Data Operations Support the Data Management Lifecycle

8.2.1. Data Operations ensure data sourcing and procurement are performed in alignment with business requirements

Data procurement should be consistent and in alignment with Data Management policies and standards as well as with the policies and standards of other business control functions (i.e.: vendor management, legal, compliance, etc.)

Capability Objectives

- Data lifecycle management is performed consistently across the organization.
- Data lifecycle is aligned with and accountable to Enterprise Data Management organization policy and standards.

8.2.2. Critical end-to-end data flows and essential attributes for in-scope business processes are defined and mapped.

Data Operations is responsible for the stewardship of the Critical Data Elements (CDEs) and the how these elements flow (lineage) across linked processes. This must be done in coordination with Enterprise Data Management office and aligned to EDM policy and standards.

Capability Objectives

- CDE and lineage have been identified and mapped across business lines in coordination with the Enterprise Data Management office
- CDE and lineage mappings are in alignment with EDM Policy and Standards.

8.2.3. The compounding processes and calculations for derived and transformed data are identified, documented and mapped

Capability Objectives

- Data transformation processes and calculations have been identified and documented across business functions in coordination with Enterprise Data Management Office.
- Data Transformation Processes are in alignment with data management policy and standards.

8.3. Data Management is aligned with cross-organizational Control Functions.

8.3.1. Control Function policies and standards reflect interdependencies with Data Management policies and standards

Data Management controls and best practices are formally included in cross-organizational control function policies and standards to ensure collaboration and alignment

Capability Objectives

- Cross-organizational references are formally included in each enterprise policy and standards
- Control teams are held accountable and subject to Corporate Audit to ensure formal coordination of each groups policy and standards**.**

8.3.2. Regular routines are established with cross-organizational control teams.

Cross-organizational teams meet regularly to keep abreast of evolving issues related to data and data operations.

Capability Objectives

- Enterprise Control functions formally coordinates with Enterprise Data Management via regular engagements, meetings and routines.

8.3.3. All data entered into the Information Ecosystem is subject to cross-organizational controls.

All new data introduced into, or delivered out of the information ecosystem is subject to cross-organizational control standards to ensure enterprise-wide compliance.

Capability Objectives

- Data introduced into or delivered out of the ecosystem is subject to design review and approval.
- Data introduced into or delivered out of the ecosystem is subject to all cross-organizational data control policy and standards.