



DAMA-DMBOK2

Data Management Body of Knowledge

Key Insights for Data Management Professionals

Based on DAMA-DMBOK2, Second Edition | DAMA International

Presentation Overview



Foundation & Strategy

Slides 3–8



Data Handling Ethics

Slides 9–10



Data Governance

Slides 11–14



Data Architecture & Modeling

Slides 15–18



Storage, Security & Integration

Slides 19–24



Master Data, DW/BI & Metadata

Slides 25–30



Data Quality & Data Science

Slides 31–35



Maturity, Organization & Change

Slides 36–40



What Is Data Management?

Chapter 1 — Foundations

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

Data as an Organizational Asset

Data is widely recognized as an enterprise asset — but managing it requires understanding its unique properties and challenges.



Not consumed when used

Unlike financial assets, data can be used simultaneously by multiple people without being depleted.



Durable but perishable in value

Data doesn't wear out physically, yet its value often changes as it ages or context shifts.



Easy to copy, hard to replace

Data is simple to duplicate and transport but extremely costly to reproduce if lost or destroyed.



Dynamic and multi-purpose

The same data can serve multiple purposes — operations, analytics, compliance, and monetization.

Data Management Principles



Unique Properties

Data differs from other assets and requires specialized management approaches.



Economic Value

The value of data can and should be expressed in measurable economic terms.



Quality Is Central

Managing data means managing its quality — fitness for purpose is a primary goal.



Metadata Required

It takes Metadata to manage data — definitions, lineage, and context are essential.



Planning Needed

Complex landscapes require architectural and process-level planning.



Cross-functional

Data management requires a range of skills, collaboration across business and IT.



Enterprise Perspective

Must be applied across the organization to be maximally effective.



Multiple Perspectives

Must account for varied viewpoints — data creators, consumers, and regulators.



Lifecycle Management

Data has a lifecycle; managing data means managing that lifecycle.



Varied Data Types

Different types of data have different management requirements.



Risk Management

Data represents risk as well as value — risks must be managed.



Leadership Commitment

Effective data management requires vision, commitment, and executive support.

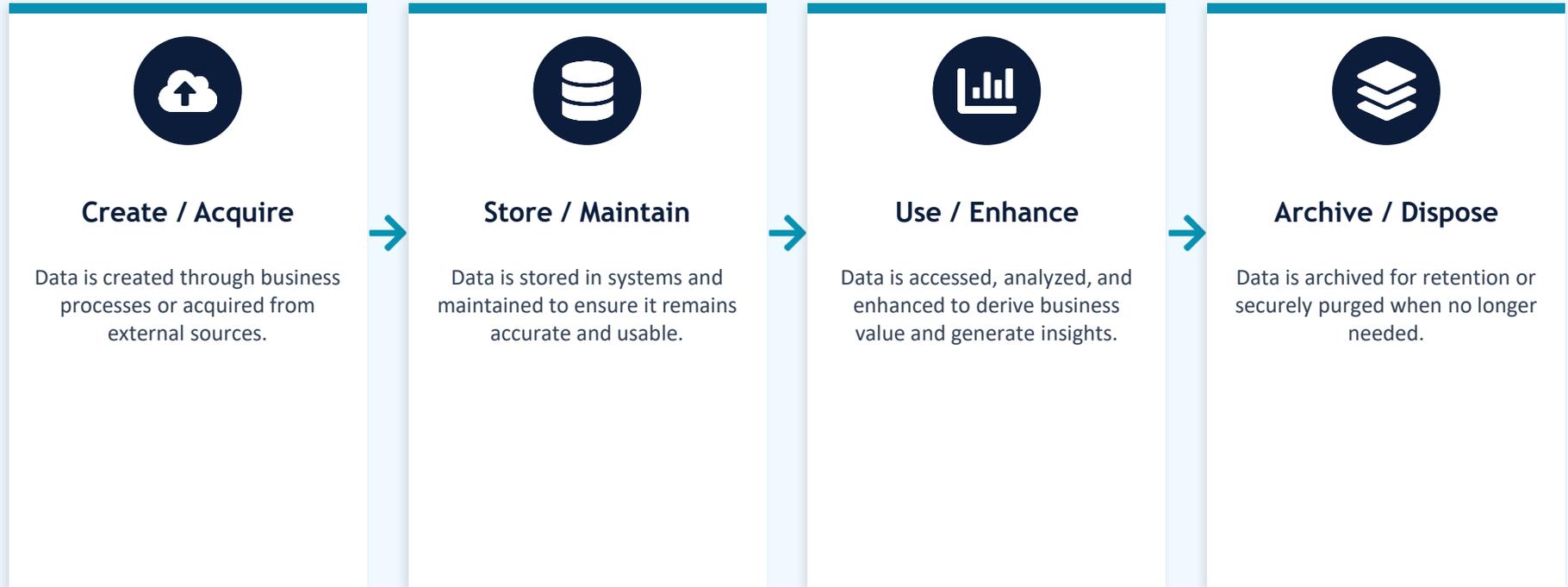
The DAMA-DMBOK2 Framework (The DAMA Wheel)

Data Governance sits at the center, surrounded by 10 Knowledge Areas that form a mature data management function.



The Data Lifecycle

Data has a lifecycle from creation to disposal. Managing data requires managing this lifecycle strategically.



Data Management Strategy Components

A data strategy should include business plans to use information to competitive advantage.

Strategy Content

-  Compelling vision for data management
-  Business case with selected examples
-  Guiding principles, values, and perspectives
-  Mission and long-term directional goals
-  Proposed measures of success

Key Deliverables

-  SMART short-term program objectives
-  Roles, organizations, and decision rights
-  Prioritized program of work
-  Draft implementation roadmap
-  Data Management Charter and Scope



Data Handling Ethics

Chapter 2 — Principles for Responsible Data Use

Ethical Principles for Data

Data ethics are principles of behavior based on ideas of right and wrong — fairness, respect, integrity, transparency, and trust.



Respect for Persons

People must be treated with dignity and autonomy. Personal data affects lives — employment, community, and freedom of choice.



Beneficence

Do not harm; maximize possible benefits and minimize possible harms. Data processing must be transparent and non-invasive.



Justice

Fair and equitable treatment. Algorithms must not disproportionately harm groups; bias in data sets must be actively countered.

Key Risks: Misleading visualizations, biased sampling, unethical aggregation, lack of lineage, and poor-quality data all carry ethical consequences. Organizations must establish an ethical data culture with governance oversight, codes of ethics, training, and whistleblower protection.



Data Governance

Chapter 3 — Authority and Control Over Data Assets

Data Governance – Definition & Scope

Data Governance is the exercise of authority and control (planning, monitoring, enforcement) over the management of data assets.



Strategy

Driving execution of Data Strategy



Policy

Setting and enforcing data policies



Standards

Data Quality and Architecture standards



Oversight

Stewardship, audit, and correction



Compliance

Meeting regulatory requirements



Issue Mgmt

Identifying and resolving data issues



Projects

Sponsoring DM improvement projects



Valuation

Defining business value of data assets

Data Stewardship – Roles & Responsibilities

Chief Data Stewards

Chair data governance bodies; may act as CDO in distributed DG organizations.

Executive Data Stewards

Senior managers serving on the Data Governance Council, providing strategic direction.

Business Data Stewards

Subject matter experts accountable for data within their domain, defining and controlling data.

Technical Data Stewards

IT professionals operating within Knowledge Areas — DBAs, BI specialists, DQ analysts.

Coordinating Stewards

Data Owners

"The best Data Stewards are often found, not made." People already stewarding data informally can be empowered through formal programs.

DG Operating Model Types

Centralized

One DG organization oversees all activities across all subject areas. Provides maximum consistency.

Replicated

Same DG model and standards adopted independently by each business unit. Promotes standardization.

Federated

One DG organization coordinates with multiple BUs to maintain consistent definitions and standards.

Success factors: Leadership and strategy, business-driven approach, shared responsibility, multi-layered governance, framework-based operations, and principle-based decision making.



Data Architecture

Chapter 4 — Blueprint for Managing Data Assets

Data Architecture – Core Components



Enterprise Data Model (EDM)

A holistic, implementation-independent conceptual or logical data model providing a consistent view of data across the enterprise. Contains key entities, relationships, guiding business rules, and critical attributes. Built incrementally through subject area models.



Data Flow Design

Defines the master blueprint for storage and processing across databases, applications, platforms, and networks. Maps data movement to business processes, locations, and technical components. Documented via matrices and data flow diagrams.



Architecture Domains

Business Architecture (capabilities, processes), Application Architecture (apps, interfaces), Technical Architecture (platforms, infrastructure), and Data Architecture (models, data flows). All domains must collaborate as each influences and constrains others.



Architecture Frameworks

Zachman Framework (6x6 ontology), TOGAF, and others provide structured approaches to enterprise architecture. They enable non-architects to understand relationships and guide strategic decisions.

Data Modeling & Design – Schemes and Levels

Data modeling discovers, analyzes, and communicates data requirements in a precise visual form called a data model.

Six Modeling Schemes

Relational

Dimensional

Object-Oriented

Fact-Based

Time-Based

NoSQL

Three Levels of Detail

Conceptual

High-level business entities and relationships. Captures scope and key terminology.

Logical

Fully attributed, implementation-independent. Detailed business rules and requirements.

Normalization – Organizing Business Complexity

Normalization applies rules to eliminate redundancy and inconsistency. Each level sorts attributes by their relationship to primary keys:

1NF: Atomic attributes, no repeating groups

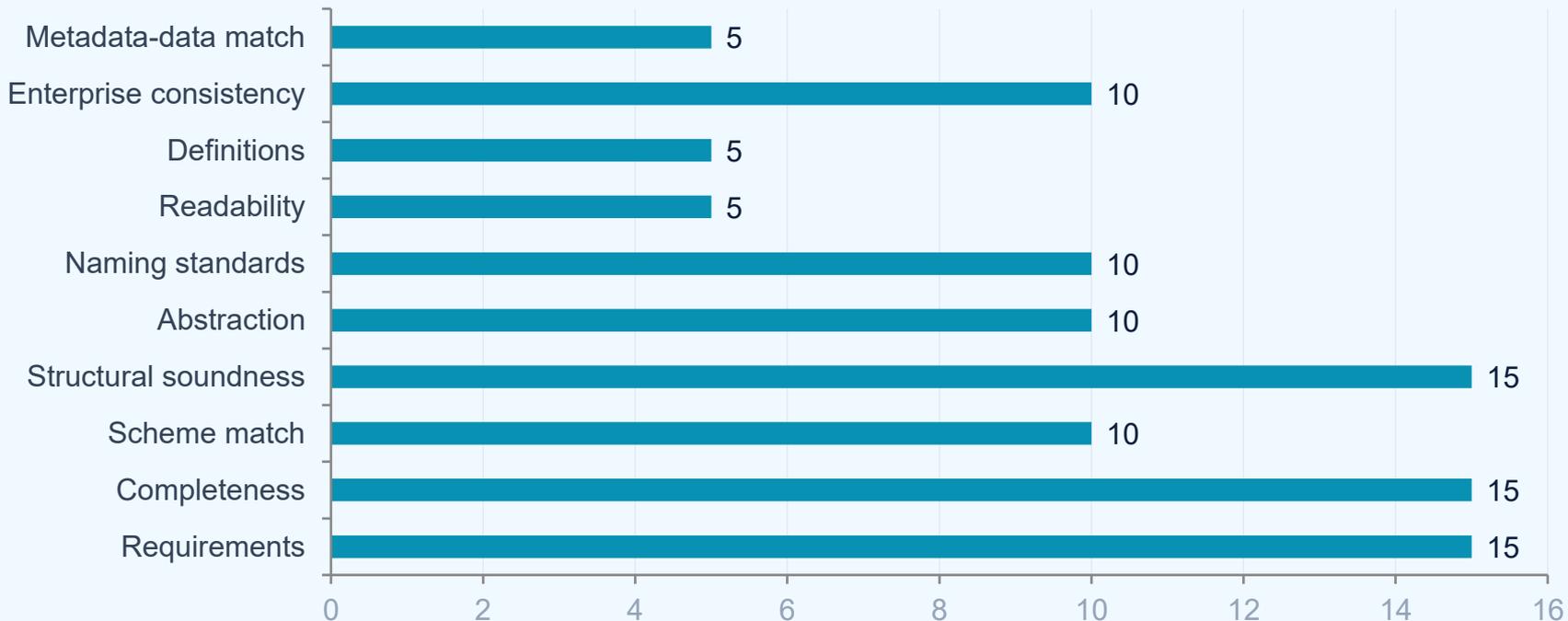
2NF: Every attribute depends on the complete primary key

3NF: No attributes depend on non-key attributes — "the key, the whole key, and nothing but the key"

BCNF/4NF/5NF: Resolve overlapping keys and multi-value dependencies (rare cases)

Data Model Scorecard® – Quality Metrics

The Data Model Scorecard provides 10 quality categories for evaluating any data model.





Data Storage & Operations

Chapter 6 — Managing Stored Data Throughout Its Lifecycle

Data Storage – Key Concepts



Database Types Spectrum

From centralized RDBMS to distributed NoSQL, columnar stores, graph databases, and cloud-native solutions. The CAP theorem governs trade-offs between Consistency, Availability, and Partition tolerance.



ACID vs BASE

Traditional RDBMS guarantees ACID properties (Atomicity, Consistency, Isolation, Durability). NoSQL systems often use BASE (Basically Available, Soft-state, Eventually consistent) for scale.



Operations Priorities

Business continuity is the primary driver. Activities include monitoring, tuning, backup/recovery, capacity planning, and ensuring SLAs for system and database performance.



DBA Role

The most established data professional role. DBAs play dominant roles in data storage operations, security, and performance. Practices are among the most mature in data management.

Guiding principle: Identify automation opportunities, build with reuse in mind (PRISM: Performance, Reusability, Integrity, Security, Maintainability)

Data Security – Protecting Data Assets

Chapter 7 — Ensuring data privacy and confidentiality are maintained, data is not breached, and access is appropriate.



Authentication & Authorization

Identity management, role hierarchies, CRUD matrices for access control.



Encryption & Masking

Data masking, encryption at rest and in transit, HTTPS, document sanitization.



Monitoring & Auditing

Intrusion detection, immediate security patch deployment, metadata tracking.



Cloud & Outsourced Security

Managing data security in cloud environments and outsourced operations.



Regulatory Compliance

GDPR, HIPAA, PCI-DSS, Sarbanes-Oxley — meeting diverse regulatory frameworks.



Classification & Marking

Data classification by sensitivity level drives security controls throughout the lifecycle.



Data Integration & Interoperability

Chapter 8 — Moving and Consolidating Data

Data Integration – ETL, ELT & Beyond



ETL Process Flow

Extract data from source systems, Transform it according to business rules and quality standards, then Load into the target system. The traditional approach for data warehousing.



ELT Process Flow

Extract and Load data into a target system first (often a data lake), then Transform. Leverages the power of modern distributed computing for transformation.



Enterprise Service Bus

Middleware architecture enabling applications to communicate through a central bus, reducing point-to-point integration complexity.



Data Virtualization

Provides a unified data access layer without physically moving data. Enables real-time access across disparate sources.



Governance of DII

Data sharing agreements, lineage documentation, and integration metrics are essential for governing data integration.

Document & Content Management

Chapter 9 — Managing the lifecycle of unstructured data and information, especially for legal and regulatory compliance.



Lifecycle Management

Plan, manage, publish, and deliver content across its lifecycle — from creation to archival and destruction.



Controlled Vocabulary & Taxonomy

Enterprise taxonomies and standard markup formats enable consistent classification, search, and retrieval.



E-Discovery

Litigation response playbooks and data maps support electronic discovery requirements and legal compliance.



Information Governance Frameworks

ANSI-859 document hierarchy, ISO 9001, and ARMA frameworks provide structure for governing content quality.



Content Governance

Control the proliferation of information, govern for quality content, and measure effectiveness through defined metrics.

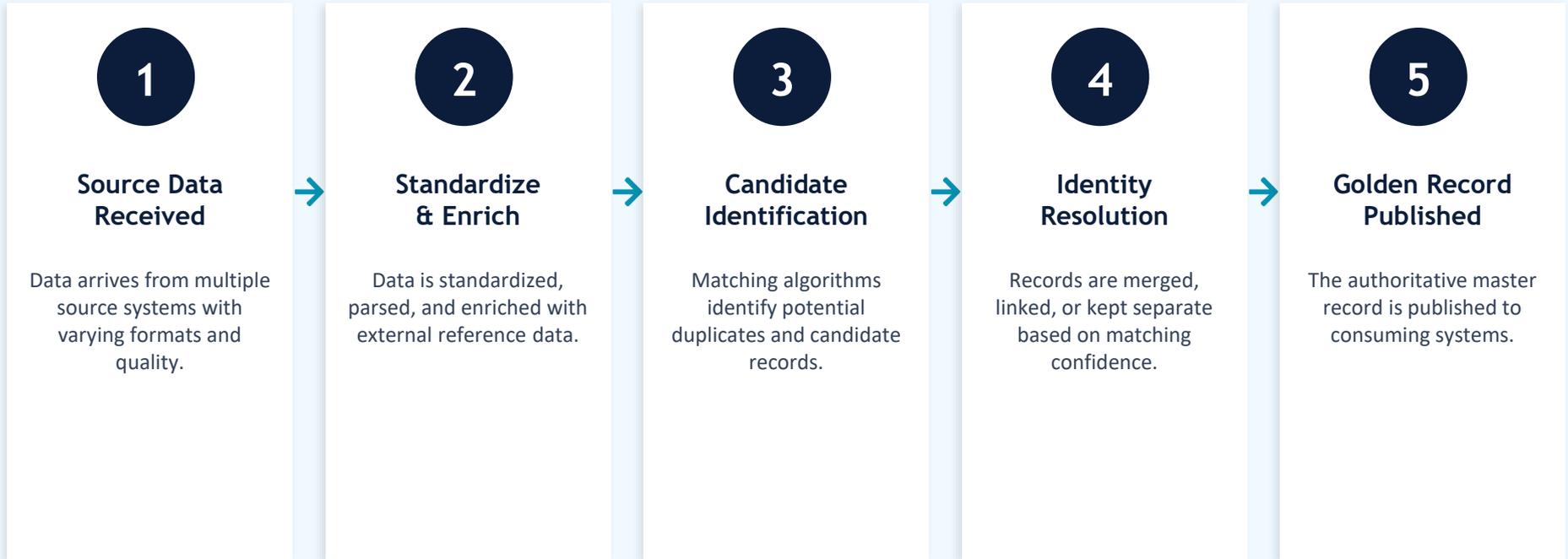


Reference & Master Data

Chapter 10 — Core Critical Shared Data

Master Data Management – Key Processing Steps

MDM maintains the most accurate, timely, and relevant version of truth about essential business entities.





Data Warehousing & Business Intelligence

Chapter 11 — Decision Support Data

DW/BI – Architecture & Key Concepts



Corporate Information Factory (Inmon)

Enterprise data warehouse with integrated, subject-oriented, non-volatile, time-variant data. Feeds dependent data marts.



Dimensional Bus (Kimball)

Bottom-up approach using conformed dimensions and facts across data marts. Bus matrix ensures enterprise consistency.



Self-Service BI

Empowering knowledge workers to create their own reports and analyses. Requires governance to maintain data quality and consistency.



CDC Techniques

Change Data Capture methods: timestamps, version numbers, status indicators, triggers, and transaction log scanning. Critical for populating the warehouse.

Key governance areas: Business acceptance, user satisfaction, SLAs, reporting strategy, and metrics for measuring DW/BI value.

Metadata Management

Chapter 12 — Planning, implementation, and control activities to enable access to high-quality, integrated Metadata.

Centralized

Single metadata repository serves the entire organization. Simplest to manage but hardest to populate.

Distributed

Multiple repositories, each owned by different teams. Flexible but risks inconsistency.

Hybrid

Central catalog with distributed capture. Balances consistency with practical team ownership.

Essential Metadata Activities

- Define Metadata Strategy
- Understand Requirements
- Define Architecture
- Create & Maintain Metadata
- Query, Report & Analyze

Data Lineage & Impact Analysis

Data lineage traces the origin and movement of data through systems. Impact analysis determines downstream effects of changes. Together, they form a critical foundation for data governance, quality, and compliance. Lineage is documented at element, system, and business rule levels.

Types of Metadata

Metadata describes what data an organization has, what it represents, how it's classified, where it came from, how it moves, and who can use it.



Business Metadata

Definitions, business rules, valid values, data ownership, stewardship assignments, quality expectations, and glossary terms. The bridge between business meaning and technical implementation.



Technical Metadata

Physical database structures, table/column definitions, data types, indexes, ETL job specifications, data lineage mappings, and system catalog information.



Operational Metadata

Job execution logs, data load statistics, error rates, processing timestamps, batch schedules, and system performance metrics. Essential for monitoring and troubleshooting.

"It takes Metadata to manage data." Metadata is itself a form of data that must be managed with the same discipline as other data assets.



Data Quality

Chapter 13 — Fitness for Use

Data Quality Dimensions

High-quality data is available, relevant, complete, accurate, consistent, timely, usable, meaningful, and understood.

Accuracy

Data correctly represents real-world entities and events

Completeness

All required data is present and no values are missing

Consistency

Data values are the same across all systems and contexts

Timeliness

Data is available when needed and reflects current state

Validity

Data conforms to syntax and business rules of its domain

Uniqueness

No duplicate records exist for the same real-world entity

Cost of Poor Data Quality: Organizations spend 10-30% of revenue on data quality issues. IBM estimated \$3.1 trillion in the US alone (2016).

Data Quality Improvement Cycle





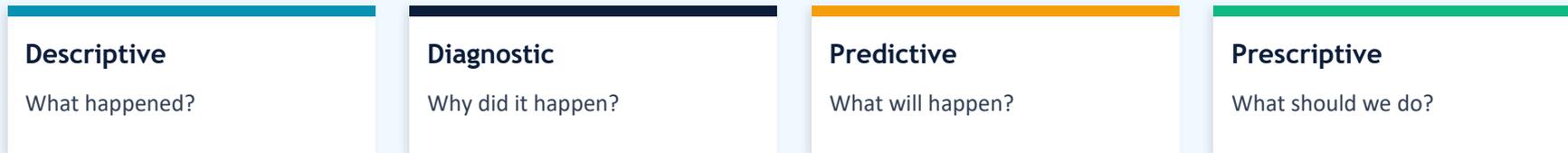
Big Data & Data Science

Chapter 14 — Emerging Capabilities

Big Data & Data Science – Process & Governance

Data Science applies quantitative and computational methods to explore data, develop hypotheses, and deploy predictive models.

Analytics Progression



Data Science Process



Big Data Governance: Visualization channel management, data science and visualization standards, data security controls for large datasets, metadata management for unstructured data, quality monitoring for streaming sources, and metrics for measuring analytical impact.

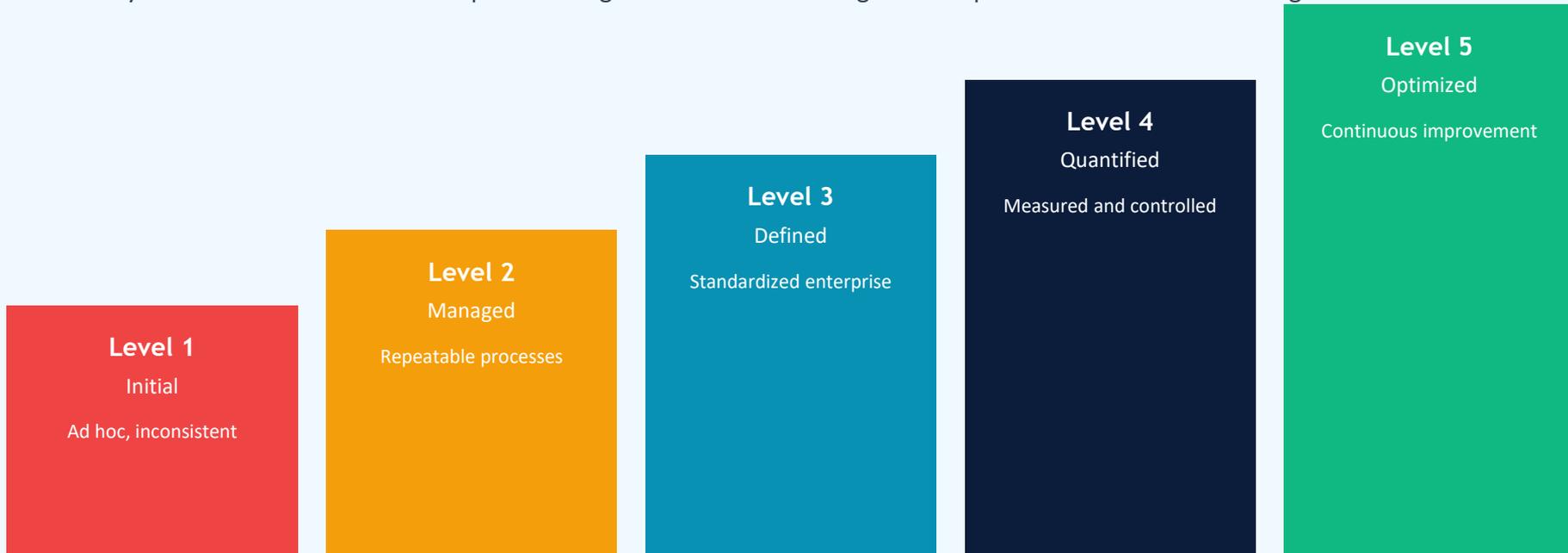


Data Management Maturity Assessment

Chapter 15 — Measuring & Improving Capabilities

DMMA – Levels & Assessment Approach

Maturity assessments evaluate and improve an organization's data management capabilities across all Knowledge Areas.



Activities: Plan Assessment → Perform Assessment → Interpret Results → Create Targeted Program → Re-Assess

Data Management Organization & Operating Models

Chapter 16 — Identifying the best organizational construct for managing data effectively.

Decentralized

Each BU manages data independently.
Maximum autonomy, risk of silos.

Network

BUs connected through informal
collaboration. Flexible but unstructured.

Centralized

Single team manages all enterprise data.
Maximum consistency.

Hybrid

Central standards with distributed execution. Balances control and
flexibility.

Federated

Central coordination with BU ownership. Best of both worlds.

Critical Success Factors: Executive sponsorship, clear vision, proactive change management, leadership alignment, stakeholder engagement, and evolution not revolution.

Organizational Change Management

Chapter 17 — Successfully embedding effective data management requires managing cultural transitions.

Kotter's 8-Stage Process for Major Change



Key Insight: "Organizational change management is more than just the 'people side of projects.' It should be viewed as how the whole organization uses to manage change well." Data management success is fundamentally about changing culture.



Assessing Your Data Management Maturity

A practical toolkit: frameworks, scorecards, metrics, and action plans to chart your path forward

The Data Management Maturity Assessment (DMMA) Framework

A structured approach to evaluating an organization's data management capabilities across all 11 DMBOK Knowledge Areas.

Why Assess Maturity?

- Establish a factual baseline — replace assumptions with evidence
- Identify highest-impact investment areas and quick wins
- Align data initiatives with business strategy and risk tolerance
- Benchmark progress year-over-year and against industry peers
- Build the executive business case for data management funding

Assessment Process in 4 Phases



1. Scope & Plan

Define KAs to assess, assemble team, set timeline & stakeholders



2. Evaluate

Score each KA using 5-level CMM scale via surveys & interviews



3. Analyze Gaps

Compare current vs. target state; prioritize by impact & feasibility



4. Action Plan

Define initiatives, owners, timelines, KPIs & quick wins

Maturity Scoring Model: The 5-Level Scale

Each Knowledge Area is assessed on a 1–5 Capability Maturity Model scale with clear, observable criteria.

1

Initial

Ad hoc, reactive, hero-dependent

- No documented processes
- No defined roles or ownership
- Data issues discovered by accident
- No metrics or monitoring

2

Managed

Repeatable within teams

- Some processes documented
- Basic roles identified informally
- Reactive issue resolution
- Limited metrics per project

3

Defined

Standardized across enterprise

- Enterprise standards published
- Formal stewardship roles assigned
- Proactive monitoring in place
- Consistent metrics & dashboards

4

Quantified

Measured & predictable

- SLAs and KPIs tracked actively
- Root cause analysis routine
- Predictive quality management
- Automated compliance checks

5

Optimized

Continuous improvement culture

- Self-correcting data ecosystems
- AI/ML-driven data management
- Industry benchmarking & leadership
- Innovation embedded in process

Target: Most organizations should aim for Level 3 enterprise-wide, with Level 4–5 for mission-critical domains.

Data Management Maturity Scorecard

Sample assessment form — score each Knowledge Area from 1 (Initial) to 5 (Optimized). Assess Current and Target state.

#	Knowledge Area	Current (1-5)	Target (1-5)	Gap	Priority	Owner
1	Data Governance	<input type="text"/>				
2	Data Architecture	<input type="text"/>				
3	Data Modeling & Design	<input type="text"/>				
4	Data Storage & Operations	<input type="text"/>				
5	Data Security	<input type="text"/>				
6	Data Integration & Interop.	<input type="text"/>				
7	Document & Content Mgmt.	<input type="text"/>				
8	Reference & Master Data	<input type="text"/>				
9	Data Warehousing & BI	<input type="text"/>				
10	Metadata Management	<input type="text"/>				
11	Data Quality	<input type="text"/>				

Instructions: Score each Knowledge Area from 1 (Initial) to 5 (Optimized) per level definitions. Gap = Target - Current. Priority = H (High), M (Medium), L (Low). Assign an executive Owner to each area.

Visual Gap Analysis: Current vs. Target Maturity

Illustrative example — a typical organization at early maturity stages with strategic targets for 18–24 months.



Key Gaps

Data Quality

Gap: +2.5 — Highest impact

MDM

Gap: +2.5 — Foundation gap

Security

Gap: +2.0 — Risk critical

Integration

Gap: +2.0 — Enabler gap

Metadata

Gap: +2.0 — Trust builder

Key Metrics & KPIs by Maturity Level

Track measurable indicators at each level to prove progress and sustain investment.

Level 1 → 2

Key Metrics

- % of critical data assets documented
- # of data owners formally assigned
- # of data issues logged (vs. discovered ad hoc)
- Time-to-resolve critical data incidents

Level 2 → 3

Key Metrics

- % of KAs with published enterprise standards
- Data governance policy adoption rate (%)
- # of stewards trained & active
- DQ score across critical data domains (%)

Level 3 → 4

Key Metrics

- DQ SLA compliance rate (%)
- Mean time to detect data issues (MTTD)
- Metadata coverage across platforms (%)
- Data lineage completeness (%)

Level 4 → 5

Key Metrics

- Automated DQ rules / total rules (%)
- Predictive issue prevention rate (%)
- Business value attributed to data initiatives (\$)
- Industry maturity benchmark percentile

Tip: Start with 3–5 metrics that matter. Avoid metric overload — it creates reporting fatigue without driving improvement.

Action Plan: Level 1 → 2 – Building Foundations (0-6 Months)

Focus on documentation, ownership, and visibility. These are the quick wins that create momentum.



Document Critical Data Assets

Inventory your top 20 data entities (customers, products, transactions). Identify source systems, owners, and known issues. Use a simple spreadsheet — don't over-engineer.



Assign Data Owners & Stewards

Identify one business owner and one technical steward for each critical data domain. Formalize with written mandate from executive sponsor.



Establish Issue Logging Process

Create a centralized data issue log (Jira, ServiceNow, or even a shared sheet). Track: issue, domain, reporter, impact, resolution, date. Make it visible.



Define Baseline DQ Rules

Implement 5–10 automated quality checks on your most critical data (e.g., null checks, referential integrity, range validation). Publish results weekly.



Launch Awareness Campaign

Brief leadership and key BI stakeholders on data management maturity, current gaps, and the cost of inaction. Build the narrative for investment.

Success Criterion: By Month 6, you should have documented ownership for 80%+ of critical data assets and a live DQ dashboard.

Action Plan: Level 2 → 3 – Enterprise Standardization (6-18 Months)

This is the hardest transition. It requires moving from team-level practices to enterprise-wide standards.



Publish Enterprise Data Standards

Define and publish naming conventions, data classification policies, modeling standards, and security protocols. Get executive sign-off.



Build Enterprise Data Architecture

Create the Enterprise Data Model and data flow diagrams. Map integration points across source systems. Define golden source for each domain.



Formalize Governance Structure

Stand up a Data Governance Council with business & IT representation. Define charter, meeting cadence, escalation paths, and decision rights.



Operationalize DQ Monitoring

Move from periodic checks to continuous monitoring. Set DQ thresholds and SLAs. Establish root cause analysis process for recurring issues.



Implement Metadata Management

Deploy a data catalog (enterprise or open-source). Ensure all critical datasets have documented business glossary terms and technical metadata.



Train & Certify Data Stewards

Develop a formal training program. Certify stewards on governance policies, tools, and escalation. Create a stewardship community of practice.

Success Criterion: Enterprise standards adopted by 3+ BUs, active governance council, metadata catalog covering 60%+ of critical assets.

Action Plan: Level 3 → 4/5 – Measurement & Optimization (18-36 Months)

At this stage, the focus shifts from building capabilities to measuring outcomes and driving continuous improvement.

Level 4: Quantified & Predictable

- 1 Implement SLAs for data delivery, quality, and availability
- 2 Deploy automated data lineage & impact analysis tools
- 3 Establish predictive DQ monitoring with anomaly detection
- 4 Integrate DQ metrics into executive dashboards & KPIs
- 5 Conduct quarterly maturity re-assessments with trend analysis

Level 5: Optimized & Innovative

- 1 Embed AI/ML into data management workflows (auto-classification, matching)
- 2 Build self-service data marketplace for business users
- 3 Implement DataOps & MLOps for end-to-end automation
- 4 Establish industry benchmarking and thought leadership
- 5 Create innovation labs for emerging data management practices

Reality check: Only ~5% of organizations globally reach Level 5. Level 3–4 across all KAs represents world-class capability for most industries.

12-Month Data Management Maturity Roadmap

A phased approach to moving from Level 1 → 3 in the first year. Adjust timelines based on organizational complexity.

Q1: Foundation

Months 1–3

- ✓ Executive sponsor secured
- ✓ Maturity assessment completed
- ✓ Critical data assets inventoried
- ✓ Data owners assigned
- ✓ Issue logging process live
- ✓ Baseline DQ metrics published

Q2: Quick Wins

Months 4–6

- ✓ Top 10 DQ rules automated
- ✓ Data governance charter drafted
- ✓ First stewardship community
- ✓ Business glossary (V1) published
- ✓ Monthly DQ dashboard live
- ✓ First executive DQ briefing

Q3: Standards

Months 7–9

- ✓ Enterprise standards published
- ✓ Governance council operational
- ✓ Metadata catalog deployed
- ✓ Data classification framework
- ✓ Steward training program
- ✓ DQ SLAs for critical domains

Q4: Scale

Months 10–12

- ✓ Standards adopted by 3+ BUs
- ✓ Continuous DQ monitoring
- ✓ Data lineage (critical paths)
- ✓ Second maturity assessment
- ✓ Year-2 roadmap & business case
- ✓ Progress report to board



Maturity Assessment: 7 Commandments

- 1 Assess honestly — vanity scores create false confidence and misallocated investment
- 2 Start where you are — Level 1 is not failure; it's the starting point for every great data organization
- 3 Prioritize ruthlessly — you cannot fix everything at once; target the 2–3 KAs with highest business impact
- 4 Assign ownership — every Knowledge Area needs a named executive owner, not a committee
- 5 Measure relentlessly — if you can't quantify improvement, you can't sustain funding
- 6 Communicate progress — share wins early and often; executive engagement requires visible results
- 7 Reassess annually — maturity is a moving target; the assessment itself drives organizational learning
"You can't manage what you don't measure — but you also can't improve what you don't assess."



Key Takeaways

- ✔ Data is a strategic asset — manage it with the same discipline as financial and physical assets
- ✔ Data Governance is the foundation — it provides oversight, not execution, of data management
- ✔ Quality is non-negotiable — poor data costs 10-30% of revenue; high quality enables competitive advantage
- ✔ Metadata makes data comprehensible — without it, data cannot be understood or trusted
- ✔ Culture eats strategy — successful data management requires organizational change, not just tools
- ✔ Start where you are — use maturity assessments to chart a pragmatic path forward