



High level approach to Data Governance and Data Quality Improvement Strategy

Data assets are growing in volume and importance, making data governance and quality a top priority. Our goal is to establish a structured approach for a comprehensive Data Governance and Data Quality framework from scratch.

This strategy aligns with industry best practices (DAMA-DMBOK and DCAM frameworks) and addresses key focus areas including data ownership, stewardship, quality, metadata, lineage, cataloging, standards, policies, and governance operating model.

Our phased roadmap will deliver quick wins while building toward long-term maturity, ensuring data becomes a trusted strategic asset for better decision-making, efficiency, and compliance.

Key Focus Areas: Data Ownership and Stewardship



Data Owners

Business executives or managers accountable for specific data domains with decision-making authority on data definitions, access, and quality requirements.



Data Stewards

Subject matter experts appointed by Data Owners who manage day-to-day data definition, monitoring, and issue resolution tasks.



Collaboration

Both roles collaborate with IT teams and with each other across domains to ensure consistency and quality throughout the data lifecycle.

Clear data ownership and stewardship are essential for accountability. Every critical data domain should have a named Data Owner responsible for its quality and usage. For a quick win, we recommend identifying a pilot domain (such as Product data) and assigning a Data Owner and Steward immediately.

Data Quality Management

Define Quality Dimensions

Establish what "good quality" means for data across dimensions like accuracy, completeness, consistency, timeliness, and uniqueness.

Resolve Issues

Track and resolve data issues through a defined workflow, building a knowledge base of common problems and solutions.



Assess Current State

Perform data profiling on key datasets to uncover anomalies such as missing values, inconsistent codes, and duplicates.

Monitor Continuously

Implement ongoing data quality monitoring with manual checks or automated rules that trigger alerts when issues arise.

As a quick win, we recommend identifying one known pain point to address immediately. For example, implementing email format validation in the data ingestion process or cleaning up invalid customer emails can demonstrate immediate value while building momentum for broader initiatives.

Metadata Management and Data Cataloging

Business Glossary

A common vocabulary defining key business terms like "Active Customer" or "Total Sales" to ensure everyone uses the same definitions. Start with 20-30 key terms maintained by Data Stewards and validated by Data Owners.

Data Dictionary

Documentation of schemas for critical databases/tables – what each field means, data types, and allowed values. Begin with a key dataset like Customer table or Sales fact table.

Data Catalog Tool

A centralized solution where users can search for data assets, see definitions, owners, and quality metrics. Consider tools like Collibra or Alation, or leverage existing platforms like SharePoint initially.

Metadata is our guide to data. Finding the right data or understanding it is likely a challenge now. Creating a centralized data catalog or even a simple glossary will help teams save time and avoid misinterpretation. We recommend starting a catalog pilot focusing on one data domain to demonstrate value.

Data Lineage (Tracing Data Flow)



Source Systems

Original data creation points (POS, e-commerce, inventory)



Transformations

How data is processed, cleaned, and modified



Data Platform

Central repository where data is stored



Reports/Analytics

Final consumption of processed data

Data lineage traces how data flows from source systems into the data platform and ultimately to reports or analyses. This visibility is crucial for trust and impact analysis. When a report number looks wrong, lineage helps identify where the breakdown occurred.

As a quick win, we recommend producing one data flow diagram for an important report or KPI. This visual can highlight how data moves and where potential quality issues might arise, reinforcing the need for governance.



Data Standards and Policies



Data Standards

Standard definitions and formats for common data elements



Data Access Policy

Who can access what data, ensuring security and privacy



Data Retention Policy

How long data is kept and when it's archived or purged



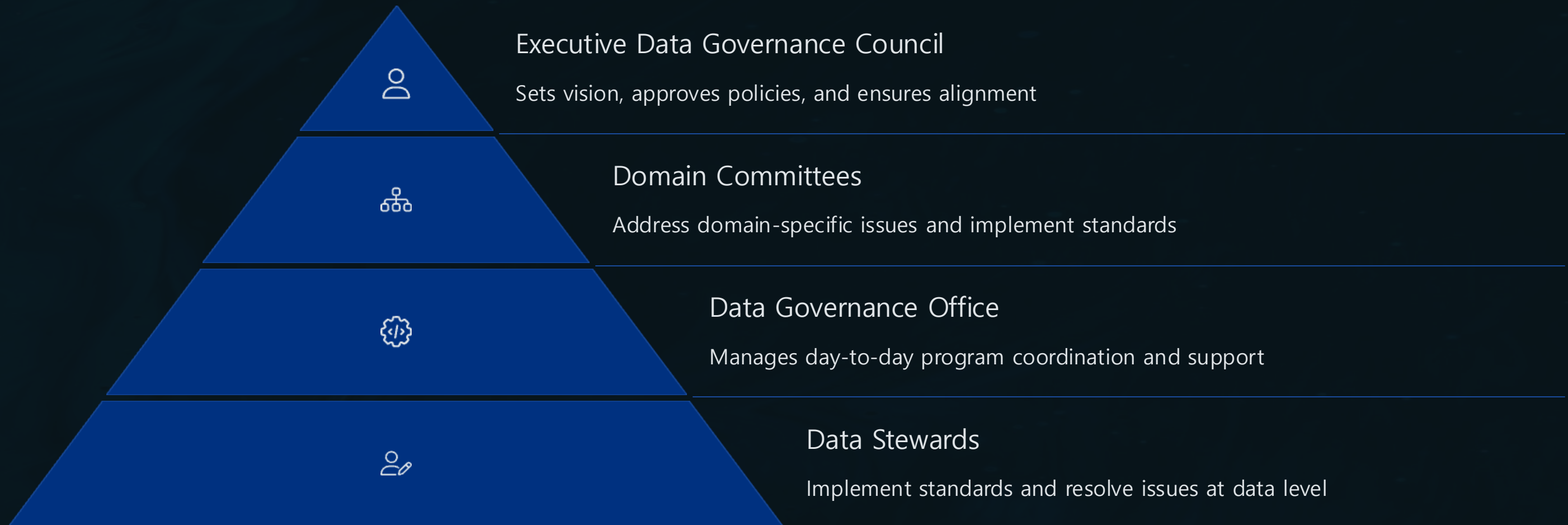
Data Quality Policy

Expectations that critical data must meet defined quality thresholds

Standards and policies ensure data is managed consistently across the organization. While they might sound bureaucratic, they're crucial for consistency and compliance. We suggest starting simple – for instance, mandating a standard date format and consistent codes across systems.

A quick win would be drafting a basic data naming and definition standard document, perhaps building on existing IT standards and addressing obvious gaps.

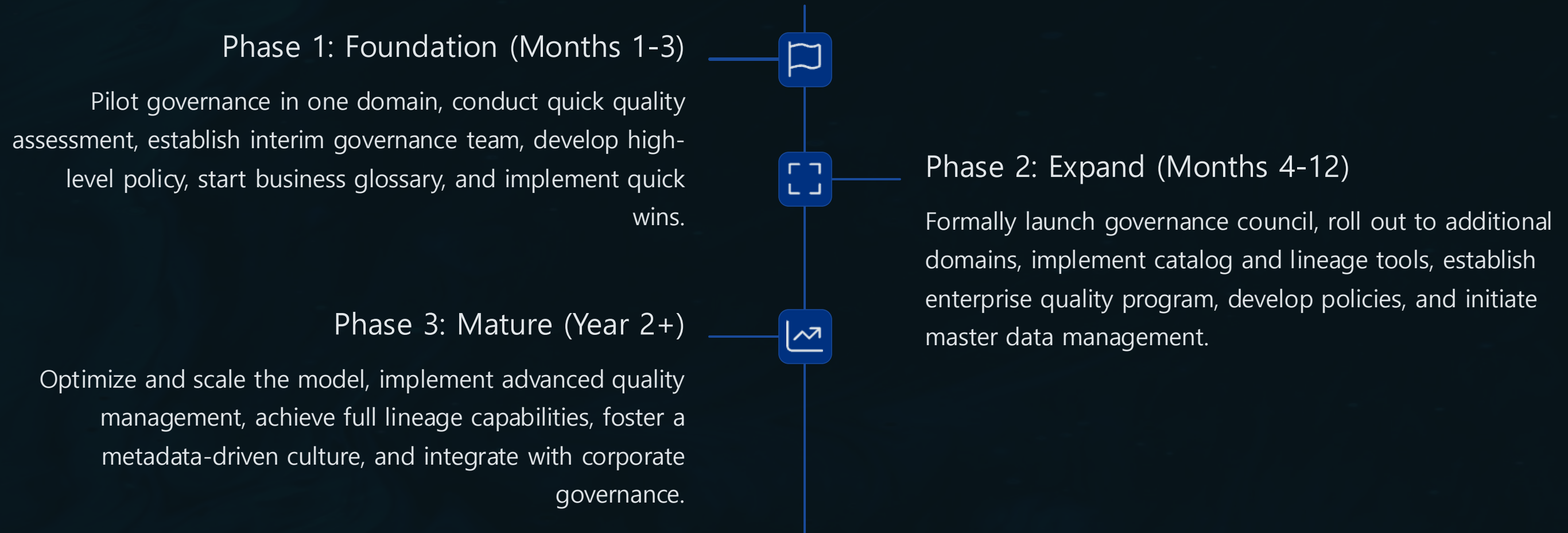
Data Governance Operating Model



To sustain governance efforts, we need the right structure. The operating model defines how we organize people and processes to govern data effectively. We propose a multi-tiered structure with clear responsibilities at each level.

As a quick win, establish a steering committee or initial version of the Data Governance Council immediately. Even if informal, this group can start meeting monthly to guide initial efforts and demonstrate the operating model at a small scale.

Phased Implementation Roadmap



This phased approach balances quick wins with strategic, long-term development. By the end of Phase 1, you should have a working prototype of data governance in one area, demonstrating value and building momentum for broader implementation.

The ultimate outcome will be a mature data governance program where data is treated as a critical asset, enabling advanced initiatives with confidence while ensuring regulatory compliance.