

National Data Management Guidelines

Master and Reference Data Management



Table of Contents

1. Master and Reference Data Management Domain Introduction

- 1.1 Domain Overview
- 1.2 Target Audience
- 1.3 Abbreviations and Glossary
- 1.4 Readability Guide
- 1.5 Stakeholders Involved
- 1.6 Key Success Factors
- 1.7 Compliance with National Data Standards
- 1.8 Domain Level Process Map

2. Master and Reference Data Management Policy

- Overview
- 2.1 MRDM Policy Creation
- 2.2 MRDM Policy Publication

3. Master and Reference Data Management Planning

- Overview
- 3.1 MRD Discovery
- 3.2 MRDM Requirements Gathering
- 3.2 MRDM Plan

4. Master and Reference Data Management Assessment and Design

- Overview
- 4.1 MRD Rules
- 4.2 MRD Architecture and Modeling

5. Master and Reference Data Management Implementation

- Overview
- 5.1 Master and Reference Data Management Tools

Table of Contents

6. Master and Reference Data Performance Management

Overview

6.1 MRDM Stewardship & Maintenance Process

6.2 MRDM KPIs

6.3 MRDM Monitoring and Reporting

6.4 QDTI

6.5 MRDM SLAs

7. Appendix

7.1 Templates and Frameworks

1. Master and Reference Data Management Domain Introduction



1.1 Domain Overview

The Master and Reference Data Management (MRDM) domain focuses on the governance, storage, and maintenance of critical data attributes, including master data (e.g., customers, products, employees) and reference data (e.g., standardized codes, classifications). The objective of this domain is to ensure the accuracy, consistency, and integrity of these data attributes across the Entity. In alignment with the National Data Standards, structured across three key dimensions—Plan, Implement, and Operate—the domain provides guidance for managing master and reference data as a critical business asset. A successful MRDM program involves the full range of data management functions, including Data Governance, Data Quality, Metadata Management, and Data Integration, ensuring effective control over data across the organization.

Master Data Management (MDM) focuses on managing master data values and identifiers to enable consistent use across systems. Its goals include ensuring the availability of accurate, timely data while reducing risks associated with ambiguous identifiers, such as duplicate or incomplete records. Reference Data Management (RDM) focuses on controlling domain-specific values and their definitions, ensuring that entities have access to a complete set of accurate and current values for each defined concept.

The purpose of the Master and Reference Data Management Guidelines is to ensure that Qatar Government Entities adhere to best practices and recommendations for maintaining complete, current, and consistent master and reference data across all processes. By following these guidelines, entities can comply with the National Master and Reference Data Management Standards, reduce the costs and risks associated with integrating data from multiple sources, minimize data redundancy, and improve data quality through standardized approaches and common data models.

Key Components of the Master and Reference Data Management Domain

Master and Reference Data Management Policy

The Master and Reference Data Management Policy is designed to provide a framework for managing master and reference data across the Entity. It formalizes the processes, roles, and responsibilities needed to ensure that this critical data is governed effectively. The policy helps align master and reference data practices with national standards and internal Entity requirements. This section covers:

- **MRDM Policy Creation:** Covers steps to create/update a master and reference data management policy, review it, and incorporate any feedback.
- **MRDM Policy Publication:** Outlines the process for conducting and auditability review, publishing the policy internally, and ensuring stakeholder awareness.

Master and Reference Data Management Planning

The planning phase identifies, prioritizes, and establishes authoritative sources for master and reference data, ensuring effective support for business operations and decision-making.

- **MRD Discovery:** Identifies and prioritizes critical data elements to ensure the most important attributes are managed accurately.
- **MRDM Requirements Gathering:** Captures overarching master and reference data management and attribute-specific requirements that support the identification and prioritization of CDEs. These requirements serve as the foundation for developing match and merge rules.
- **MRDM Plan:** Provides a roadmap for ensuring adequate governance, storage, and maintenance of master and reference data in alignment with organizational goals.

Master and Reference Data Management Assessment and Design

This phase focuses on designing master and reference data match and merge rules to enable the creation of a golden record and establish a robust architecture for managing master and reference data effectively

- **MRD Rules:** Establish the groundwork for defining attribute trust, similarity, and survivorship rules to support the match and merge processes.
- **MRD Architecture and Modeling:** Select the appropriate MDM Hub architecture to manage master and reference data. Ensure alignment between the conceptual architecture, information architecture, MDM hub technical requirements and the Entity's broader goals and requirements.

Master and Reference Data Management Implementation

This phase involves deploying tools and technologies to effectively manage master and reference data

- **Master and Reference Data Management Tools:** Select an MDM tool as the foundation for managing master and reference data and develop the MDM Hub in alignment with the MDM conceptual and information architecture and MDM hub technical requirements.

Master and Reference Data Performance Management

This phase ensures the continuous monitoring, maintenance, and optimization of master and reference data processes

- **MRDM Stewardship and Maintenance Process:** Establish ongoing stewardship roles and responsibilities to ensure the quality and consistency of master and reference data.
- **MRDM KPIs:** Track data quality, accuracy, and consistency using established KPIs to evaluate the effectiveness of management efforts.
- **MRDM Monitoring and Reporting:** Monitor and report on data performance, using reporting tools to track metrics and identify improvement areas.
- **QDTI:** Evaluating the Entity's adherence to the National Master and Reference Data Management Standards as part of the broader QDTI framework.
- **MRDM Service Level Agreements (SLAs):** Define and monitor SLAs to set expectations for data quality, availability, and timelines for issue resolution.

1.2 Target Audience

This handbook is intended for the Data Management and Security Director. It is also suitable for team leads and other individuals deemed responsible for managing master and reference data. The handbook serves as a resource, offering practical guidance to help establish the essential components of the Entity's data management program, setting the foundation for its data management journey.

1.3 Abbreviations and Glossary

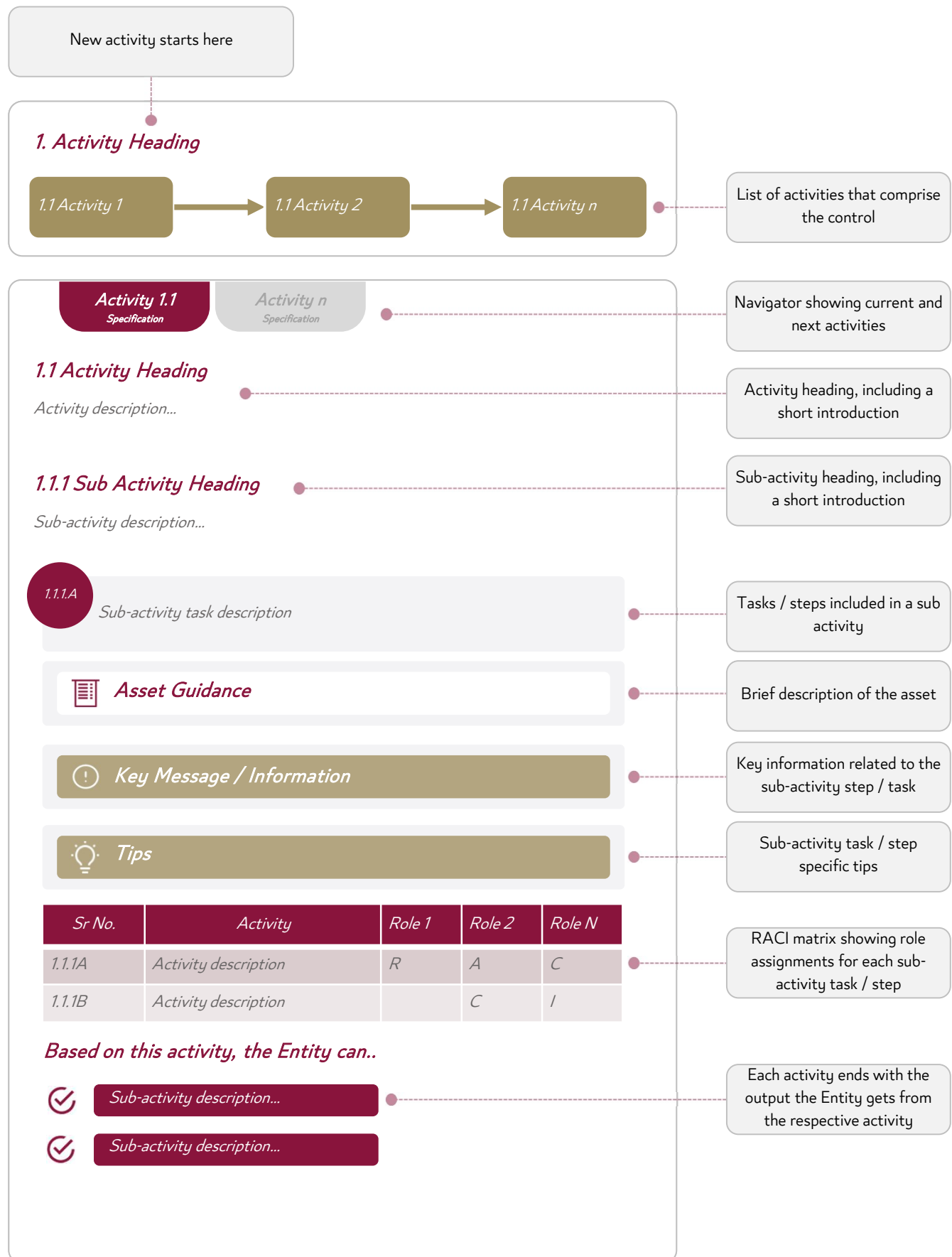
Abbreviation/Term	Definition / Long Name
API (Application Programming Interface)	A set of protocols and tools that allows different software applications to communicate and share data.
Attribute Similarity Rule	A rule that defines the methods and criteria for identifying similarities between data attributes across systems.
Attribute Trust Rule	A rule that defines the level of reliability or confidence associated with specific attributes from a data source.
Authoritative Data Source	Name of the entity owning the mandate of the dataset or attribute.
Business Data Steward	A person responsible for ensuring the accuracy and governance of data used within specific business functions.
Business Processes	A business process consists of a series of interconnected tasks or activities aimed at achieving a specific business goal or objective. These processes may involve multiple stakeholders, inputs, actions, and outputs, and can be defined at varying levels of detail. Clearly defining business processes is essential for enhancing efficiency, identifying areas for improvement, and achieving organizational goals.
CDE	A Prioritized Attribute (CDE) is data that is essential to the successful operation and decision-making processes of an Entity. The accuracy, completeness, and timeliness of CDEs are vital to the Entity's operations and strategies. These elements are often required for regulatory compliance, financial reporting, business policy adherence, or day-to-day operations. Managing CDEs effectively enhances data quality, reduces risks, and supports informed decision-making. Examples of CDEs include personal identifiable information (PII), financial transaction data, product identifiers, and vendor information.
Conceptual Architecture	The Conceptual Architecture will contain descriptions of the overall architecture components defined to support the implementation of the target Master and Reference management as per the selected hub for implementation.
Coexistence Hub	A Coexistence Hub allows master data to reside in multiple systems but synchronizes it with a central master data system. This model maintains a "golden record" in the MDM system, which is updated and synchronized with other systems as needed, ensuring data consistency while allowing distributed management.
Centralized Hub	In a Centralized Hub model, all master data is consolidated into a single, authoritative system of record. This system serves as the enterprise's trusted source for master data, ensuring uniformity and consistency across all systems and processes that rely on master data.
Cross Entity Processes	A cross-entity process is a business process that involves multiple Entities or enterprises, requiring the exchange of information, products, or services among different stakeholders. Examples of cross-entity processes include financial transactions, supply chain management, and logistics operations.
CRUD (Create, Read, Update, Delete)	The four basic operations performed on data within a system.
DAMA	Data Management Association
Data Catalog	A data catalog is a centralized inventory of all data assets within an Entity. It provides a comprehensive view of an Entity's data assets, their metadata, lineage, and data usage.

Abbreviation/Term	Definition / Long Name
Data Domain	A data domain is a distinct, well-defined subset of an Entity's data that shares common characteristics, such as its nature or the systems that store it. Data domains play a crucial role in effective data management by enabling organizations to define data ownership, establish governance policies, and implement data quality controls tailored to each domain.
Data Model	A data model is a structured way of defining and organizing data for a system or organization. It shows how data is related, how it is stored, and the rules that govern it. Data models help create a common understanding of data between business and IT, ensuring consistency, improving data quality, and guiding how systems manage and use data.
DGO	National Data Governance Office
DM	Data Management
DMBOK	Data Management Body of Knowledge
Data Organization	Data organization refers to a data sector in the case of a large-scale data management program and a data office for small and medium scale data management programs
External Reference Data	External reference data signifies the reference data that is obtained from sources outside of an Entity's internal systems and databases. This data can come from a variety of sources, such as external data providers, public records, third party data sources etc.
Golden Record	A golden record for reference and master data is the single, trusted source that represents the most accurate, complete, and authoritative version of a master or reference data entity within an Entity. It is typically created by consolidating, cleansing, and reconciling data from various source systems to provide a unified view. The golden record is managed through Master Data Management (MDM) processes, with data stewards or data management teams responsible for ensuring its ongoing accuracy, consistency, and governance across the enterprise.
Internal Reference Data	The reference data that is owned and managed within the Entity is termed as Internal Reference Data.
Information Architecture	Information Architecture refers to the structured design of data, components, and processes within the MDM hub. It defines how master and reference data is ingested, validated, stored, and distributed, aligning with governance policies and business objectives.
LDAP (Lightweight Directory Access Protocol)	A protocol used to manage and access distributed directory information.
Master Data	Master data refers to the core data that provides the context for business activity in the form of common and abstract concepts that relate to the activity. It includes the details of (definition and identifier) internal and external objects involved in the business transactions, such as customers, products, employees, vendors and controlled domains (code values)
MRD	Master and Reference Data
MRDM	Master and Reference Data Management

Abbreviation/Term	Definition / Long Name
Master Data Objects	A master data object refers to a core business entity that is vital for multiple systems and processes within an Entity. These objects can represent key entities such as customers, products, vendors, locations, and employees, and they are consistently referenced across different business processes. A master data object typically includes attributes, identifiers, and metadata that describe these entities, and its role is to ensure consistent data representation across the enterprise.
MDM Hub	A Master and Reference Data (MRD) Hub is a system that maintains a single source of truth for master and reference data. It ensures that standardized data is consistently applied across various applications, fostering data accuracy and reliability throughout the Entity.
MDM Repository	The central component of the MDM Hub, designed to store, manage, and govern cleansed, validated, and consolidated Master and Reference Data (MRD). It serves as the single source of truth, ensuring data consistency, accuracy, and integrity across all consuming systems.
Matching and Merging Rules	Matching and merging rules define how data records that refer to the same entity are identified and combined. These rules help create a unified "golden record" by reconciling discrepancies and resolving conflicts between data from different sources.
QDKC	Qatar Data Knowledge Case
Reference Data	Reference data is a collection of values which is typically used for categorizing or classification of other assets within an Entity. This collection of values is static in nature e.g., they do not change frequently often over time. Managing this collection in a global repository is needed to achieve standardization across the Entity which in turn will help all the subsystems follow the same terminology hence aiding same understanding across the board. Reference data can be shared among multiple systems. Some commonly used reference data types are Reference List, Cross reference list, Taxonomic reference data, Industry reference data, proprietary or internal reference data, etc.
Reference Data Objects	Reference data objects are required to classify or categorize the master data objects. Reference data objects are often used as a lookup-up table or reference tables to provide context and meaning to the master data objects. Examples of Reference Data Objects are Country code, product codes, customer segment etc.
Registry Hub	A Registry Hub stores only the minimum data necessary to uniquely identify master data records, while the detailed data remains in its source systems. This hub ensures cross-referencing and consistent identification across systems without duplicating data, facilitating the coordination of master data without a full consolidation.
Repository Hub	In a Repository Hub model, master data is collected from multiple source systems, cleansed, deduplicated, and consolidated into a single location. The "golden records" created here serve as the authoritative source of truth, enabling consistent downstream reporting and operational usage.

Abbreviation/Term	Definition / Long Name
Survivorship Rule	A rule that determines which attribute values are retained when consolidating duplicate records into a single golden record.
SME	Subject Matter Expert
Technical Validation	Technical validation refers to the process of applying systematic checks and validation methods to assess attribute trust levels within defined source systems. It helps determine which attributes are deemed the most trustworthy during the match and merge process.
Validation Rules	Predefined criteria or checks to ensure data quality, such as ensuring numerical values fall within an acceptable range or text fields match a specific format.

1.4 Readability Guide



Note: Should an activity only contain one sub activity, this sub activity will be in the format of "1.1.A" instead of "1.1.1.A"

1.5 Stakeholders Involved

Below are the stakeholders and actors who contribute to the different processes outlined in the upcoming chapters.

Stakeholder	Role Description
Data Governance Office (DGO)	A Data Governance Office (DGO) is the centralized national body in the National Planning Council responsible for facilitating and coordinating data governance efforts within all associated government entities. It serves as the principal body overseeing the management of data as a strategic asset, ensuring that data-related processes align with the national objectives and comply with relevant policies and regulations.
Data Management & Security Director	The Data Management and Security Director is responsible for leading and overseeing the development and implementation of the data agenda within the government entity, ensuring alignment with strategic objectives and compliance with security regulations. This role promotes best practices in data management while ensuring robust security measures are in place to safeguard the Entity's data assets. The director bridges business and IT to support data-driven decision-making, elevate data as a strategic asset, and ensure data is managed and secured across its lifecycle. Additionally, the director fosters a culture of data security and literacy within the entity, supervises the value realization agenda for data, and ensures alignment with national security policies and standards, as well as the vision of executive leadership.
Data Management and Governance Section Head	The Data Management and Governance Section Head is responsible for leading, managing, and overseeing both data management and governance initiatives within the government entity. This role focuses on establishing and implementing best practices to ensure data governance and management align with the Entity's business requirements, strategic objectives, and compliance with national data standards. The Section Head collaborates with various teams across the Entity to develop and operationalize frameworks, policies, and processes for data governance and management. This includes improving data quality, accessibility, and protection, while also ensuring data lifecycle management practices are in place. The Section Head fosters a data-driven culture and ensures that data is treated as a strategic asset to drive decision-making and operational excellence.

Stakeholder	Role Description
Data Architecture and Analytics Section Head	The Data Architecture and Analytics Section Head oversees the data architecture and analytics agenda/initiatives within the government entity, ensuring alignment with strategic objectives and national data management standards. This role is responsible for designing and maintaining a secure, scalable data infrastructure that supports data integration, management, and accessibility, as well as Overseeing data modeling, database design, and system integration, developing advanced analytics, data visualization, and reporting solutions to enable informed decision-making. The Section Head collaborates with departments to identify key performance indicators (KPIs), analyze trends, and provide actionable insights that optimize processes, enhance service delivery, and support policy-making. By combining expertise in data architecture and analytics, this role drives a culture of data-driven innovation and operational excellence across the Entity.
Data Governance Specialist	The Data Governance Specialist oversees the frameworks, policies, and procedures of governance and ensures their implementation to guarantee the effective management and utilization of data and resources within the Entity. This role includes developing and maintaining governance structures, leading compliance initiatives, and fostering a culture of accountability and transparency.
Data Quality Specialist	The Data Quality Specialist is responsible for overseeing data quality management processes within the entity. This role involves implementing data quality standards, monitoring data quality, and ensuring data consistency and effective usage across the Entity. The Data Quality Specialist collaborates with various teams to address data issues and enhance overall data quality.
Metadata Specialist	The Metadata Specialist is responsible for managing the Entity's data catalog and metadata assets to enhance data visibility, accessibility, and governance. This role focuses on developing and maintaining metadata standards, ensuring metadata quality, and supporting integration with the data catalog. The specialist plays a key role in enabling data discovery, supporting data-driven initiatives, and ensuring alignment with the Entity's data governance and management strategies.

Stakeholder	Role Description
Data Architect	The Data Architect is responsible for designing and implementing the Entity's overall data architecture, including logical and physical data models. This involves creating secure, scalable data structures that align with business needs, security requirements, and national and international data management and governance standards. The Architect ensures proper data integration, quality, and accessibility across systems, working with teams to maintain consistency and accuracy, and support data-driven decision-making to optimize data as a strategic asset.
Data Strategy Expert	The Data Strategy Expert is tasked with developing and promoting a data management and analytics strategy to achieve broad acceptance across the government entity. The expert identifies innovative ways to utilize data to achieve strategic objectives and enhance effective decision-making. Additionally, they introduce frameworks and practices to deliver accurate, reliable, and actionable data insights.
Business Data Analyst	The Business Data Analyst is responsible for collecting and analyzing data from various sources to extract valuable insights that support strategic decision-making within the Entity. The specialist designs reports and dashboards to present analyses in clear and understandable ways and works closely with other teams to ensure data quality and alignment with business needs.
Data Visualization Officer	The Data Visualization Officer is responsible for building dashboards and visual reports based on data, which are used to provide insights to leadership for decision-making.
Data Owners	Data Owners are responsible for their department's (or sub-department's) data and its quality throughout its lifecycle across systems and ownership boundaries. They are accountable for supporting, guiding, reviewing, correcting, and/or approving the work of the data stewards managing their data to ensure their datasets are maintained to the highest possible standards.
Data Stewards	Data Stewards are responsible for identifying datasets related to relevant subdomains for inventory purposes and monitoring data quality metrics within and across entities. They are nominated as a business representative with a deep understanding of a specific data subdomain and can identify functional data requirements and needs.

1.6 Key Success Factors

Executive Sponsorship: Strong executive sponsorship is essential for the success of master and reference data management initiatives. Leadership must provide oversight, direction, and support to ensure that MRDM practices are designed, implemented, and maintained effectively. Ongoing commitment and involvement are critical to resolving challenges and ensuring alignment with the Entity's strategic goals. Ensuring that executive sponsors prioritize shared ownership and accountability for master and reference data further strengthens governance and encourages cross-departmental collaboration.

Engaged Stakeholders: The active involvement of stakeholders—including data stewards, IT teams, and business owners—is essential for the success of MRDM efforts. Engaged stakeholders ensure that master and reference data management addresses practical needs, encourages adoption, and fosters alignment across departments. Involving stakeholders early in the process also reinforces principles such as ownership and stewardship, ensuring that master and reference data are managed and shared consistently across the Entity.

Realistic Implementation Roadmap: A practical and detailed implementation roadmap is necessary to guide MRDM initiatives. This roadmap should account for technical dependencies, resource availability, time constraints, and organizational readiness to ensure achievable milestones and objectives.

Focus on Data Quality and Integrity: Maintaining high data quality is a cornerstone of MRDM success. This requires implementing robust validation, matching, and survivorship rules to eliminate duplicates, ensure accuracy, and create reliable golden records for master and reference data. Ongoing data quality monitoring should be integrated into MRDM to maintain consistency and ensure adherence to the national data quality management.

Robust Architecture and Tools: Selecting the right MRDM architecture and tools is crucial to achieving accuracy, consistency, and efficiency. The Entity must ensure that the tools implemented support foundational, advanced, and very advanced capabilities, such as workflow management, validation, and AI-powered enhancements, aligned with business needs and scalability requirements. Adopting a common MDM hub architecture ensures alignment with enterprise data models and reduces complexity across Entity processes.

Continuous Improvement: Regularly reviewing and refining MRDM processes and tools ensures they remain relevant and effective. Feedback from users, advancements in technology, and evolving business needs should be incorporated to enhance the MRDM framework over time.

The following chapters will provide detailed guidance in each of these areas, offering actionable steps to establish robust master and reference data management practices and drive data consistency, accuracy, and operational efficiency within the Entity.

1.7 Compliance with National Data Standards

National Data Standards		QDKC
Control	Specification	Activity
MRDM.1.1 MRDM Policy	MRDM.1.1.1 MRDM Policy Creation	2.1 MRDM Policy Creation
	MRDM.1.1.2 MRDM Policy Publication	2.2 MRDM Policy Publication
MRDM.1.2 MRDM Planning	MRDM.1.2.1 MRD Discovery & Prioritization	3.1 MRD Discovery
	MRDM.1.2.3 MRDM Plan	3.3 MRDM Plan
MRDM.2.1 MRDM Assessment & Design	MRDM.2.1.1 MRD Rules	4.1 MRD Rules
	MRDM.2.1.2 MRD Architecture & Modeling	4.2 MRD Architecture & Modeling
MRDM.2.2 MRDM Implementation	MRDM.2.2.1 MDM Tools	5.1 Master and Reference Data Management Tools
	MRDM.2.2.2 MDM Tools	5.1 Master and Reference Data Management Tools
	MRDM.2.2.3 MDM Tools	5.1 Master and Reference Data Management Tools
MRDM.3.1 Master & Reference Data Performance Management	MRDM.3.1.1 MRDM Stewardship & Maintenance Process	6.1 MRDM Stewardship & Maintenance Process
	MRDM.3.1.2 MRDM KPIs	6.2 MRDM KPIs
	MRDM.3.1.3 MRDM Monitoring & Reporting	6.3 MRDM Monitoring & Reporting
	MRDM.3.1.4 MRDM Monitoring & Reporting	6.3 MRDM Monitoring & Reporting
	MRDM.3.1.5 MRDM SLAs	6.5 MRDM SLAs
	MRDM.3.1.6 MRDM SLAs	6.5 MRDM SLAs

Legend: Capability Level



Foundational

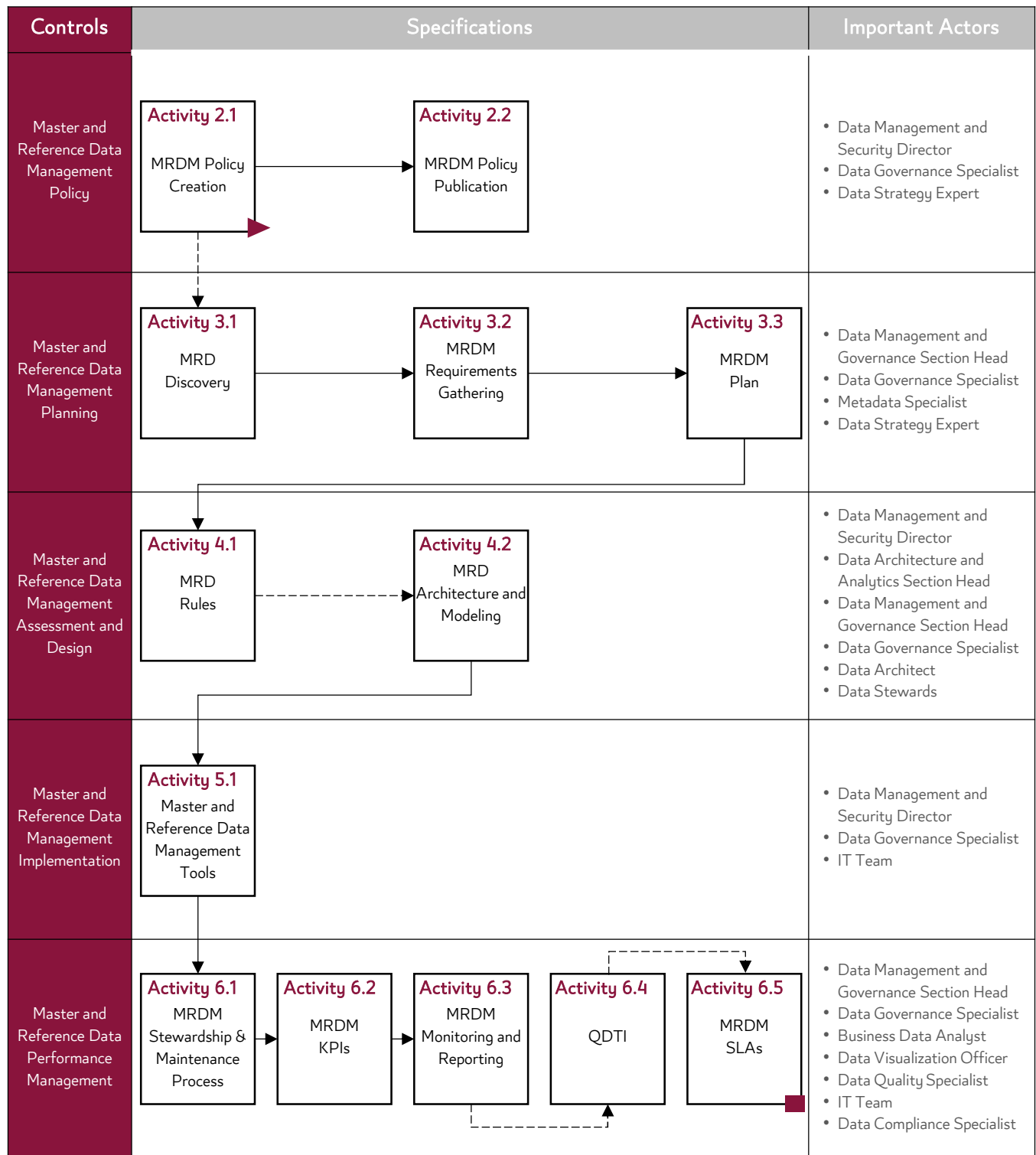


Advanced



Very Advanced

1.8 Domain Level Process Map



2. Master and Reference Data Management Policy



2. Master and Reference Data Management Policy Overview

What is a Master and Reference Data Management Policy and why does the Entity need it?

A Master and Reference Data Management (MRDM) policy defines the framework and principles for governing, storing, and maintaining an Entity's most critical data assets. It enforces best practices to ensure key data elements are consistently accurate, reliable, and controlled, serving as a single source of truth across the Entity. This policy plays a pivotal role in aligning data management efforts with Entity goals and regulatory requirements. The need for a master and reference data management policy arises from several key factors:

1. **Data Consistency:** It ensures uniformity of data across departments by establishing a single source of truth. This minimizes errors caused by conflicting versions of data, improves collaboration, and enables reliable reporting and decision-making.
2. **Data Quality:** It establishes processes to maintain high-quality data by addressing issues such as duplicates, missing values, and inconsistent classifications. This ensures the accuracy and reliability of critical data used for operations and decision-making.
3. **Data Ownership and Stewardship:** It defines clear roles and responsibilities for data ownership and stewardship, ensuring accountability for the accuracy, security, and governance of master and reference data. This helps resolve data disputes and maintain governance alignment across the organization.
4. **Compliance and Risk Management:** It helps meet regulatory requirements and mitigates risks from inaccurate or inconsistent data.
5. **Operational Efficiency:** Centralized data governance reduces redundancies and inefficiencies in data processes by centralizing data governance and enabling smoother workflows. A well-structured policy accelerates access to and usage of data, improving overall productivity.
6. **Data Integration and Interoperability:** Standardized data definitions support seamless integration across systems and applications. This reduces data silos, promotes system compatibility, and enhances collaboration across organizational units.
7. **Data Value Realization:** It maximizes the value of master and reference data by enabling its efficient use in analytics, operations, and customer engagement. This ensures that data contributes to achieving organizational objectives and driving innovation.

By implementing a robust MRDM policy, Entities can enhance decision-making, streamline operations, and foster trust in the data that drives business performance and innovation.

The remaining sections of this chapter will detail the below activities to guide the execution of the below:



Guiding Principles for Master and Reference Data Management Policy

1. **Single Source of Truth:** Maintain master and reference data as the authoritative source across the Entity to eliminate discrepancies.
2. **Data Quality:** Enforce high standards for accuracy and completeness to ensure data is fit for operational and analytical use.
3. **Consistency:** Ensure master and reference data are uniformly defined and used across all systems and departments.
4. **Ownership and Stewardship:** Assign clear responsibility for the governance and maintenance of master and reference data.
5. **Compliance and Security:** Ensure data management practices comply with legal, regulatory, and security requirements.
6. **Accessibility:** Make master and reference data easily accessible to authorized users for decision-making and operations.
7. **Governance Alignment:** Integrate master and reference data management within the broader data governance framework.
8. **Continuous Improvement:** Regularly monitor and improve data management practices to meet evolving business needs.

2.1 MRDM Policy Creation

The entity should establish a Master and Reference Data Management Policy to ensure data is managed and governed effectively, securely, and efficiently throughout its lifecycle. The policy should include, at a minimum:

- Overview, purpose, scope, and objectives
- Policy applicability, validity, and review process
- Policy guidelines aligned with the National Data Management Policy and Standards
- Detailed policy statements covering:
 - Master and reference data plan
 - Identification of master and reference data
 - Implementation of technology to support master and reference data management
 - Operation of master and reference data to support consuming systems
 - Performance management and monitoring
 - Escalation and issue management
 - Periodic policy review and updates
 - Key policy performance indicator
 - Roles and responsibilities
- References

Process – MRDM Policy Creation

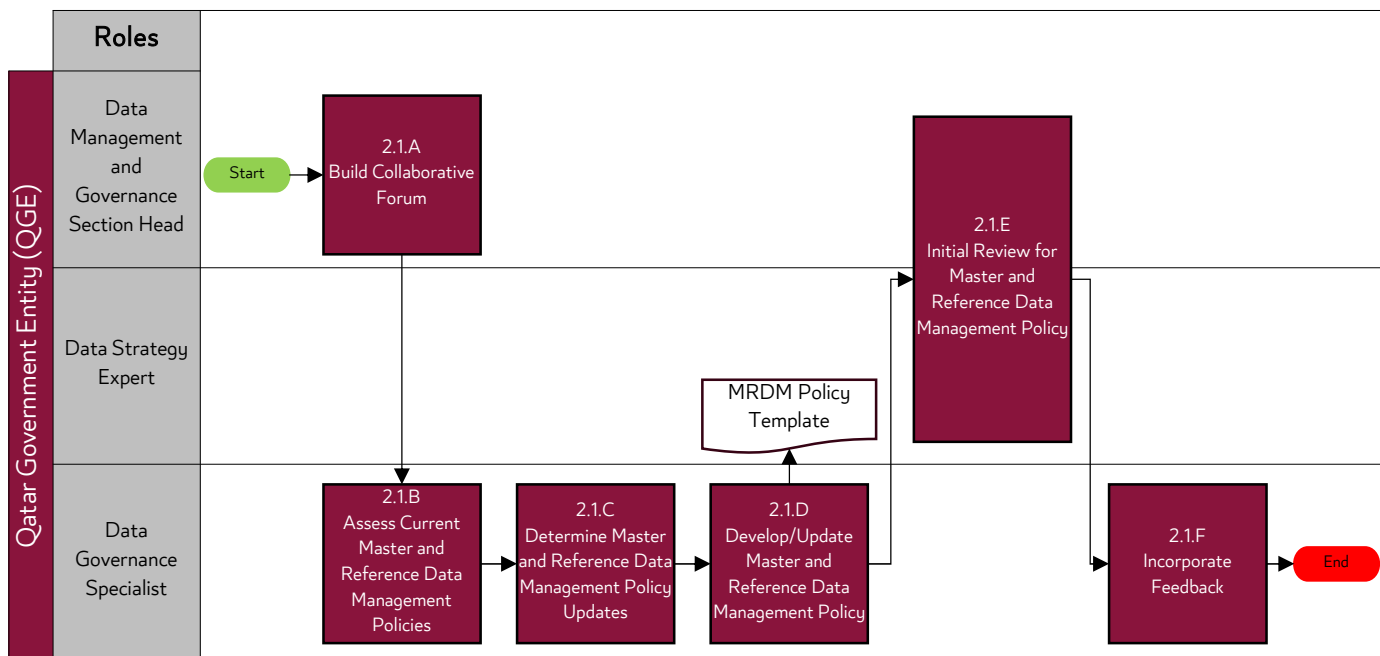


Figure 1: MRDM Policy Creation Process Map

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

2.1.A

Build Collaborative Forum: The Data Governance Specialist will be responsible for drafting the policy document. The Data Management and Governance Section Head creates a collaborative forum of subject matter experts (SMEs) whose input will be essential for the Data Governance Specialist to develop the policy. This group will ensure the policy is comprehensive and aligned with needs of the entity.

2.1.B

Assess Current Master and Reference Data Management Policies: Assess the current state of the entity's MRD policy, if one exists, to identify gaps, overlaps, and misalignment with the data management strategy. This assessment ensures the policy aligns with strategic goals and supports effective data management. When conducting the assessment, consider the following criteria:

- **Relevance to Strategic Objectives:** Evaluate whether the policy aligns with the Entity's broader strategic goals and operational priorities
- **Compliance with National Standards:** Assess the policy for compliance with relevant national data standards
- **Clarity and Accessibility:** Review the language, structure, and accessibility of the policy to ensure it is easy to understand and follow by all relevant stakeholders

2.1.C

Determine Master and Reference Data Management Policy Updates: Based on the current state assessment and inputs from relevant SMEs, define a comprehensive plan for the policy that needs to be created, updated, or realigned.

2.1.D

Develop/Update Master and Reference Data Management Policy: Using the Master and Reference Data Management Policy Template, the Data Governance Specialist develops or updates the document as a preliminary policy.

Asset Guidance

The **Master and Reference Data Management Policy Template** is a sample table of contents that the Entity can refer to when drafting its policy. It includes the topics required by the standards, as well as additional recommended topics to consider. It also gives a sample of policy statements for one topic from the table of content to help entities draft the policy document.

Illustrative

Table of Contents	
1	Introduction.....#
1.1	Overview.....#
1.2	Purpose.....#
1.3	Scope & Applicability.....#
1.4	Objectives.....#
2	Master and Reference Data Management Policies.....#
2.1	Data Governance and Integrity.....#
2.2	Master And Reference Data Plan.....#
2.3	Identification of Master and Reference Data.....#
2.4	Implementation of Technology to Support Master and Reference Data Management.....#
2.5	Operation of Master and Reference Data to Support Consuming Systems.....#
2.6	Performance Management and Monitoring.....#
2.7	Escalation and Issue Management.....#
2.8	Disaster Recovery and Business Continuity.....#
2.9	Collaboration and Continuous Improvement.....#
2.10	Change Management and Version Control.....#
2.11	Data Security and Access Control.....#
3	Periodic policy review and updates.....#
3.1	Revision Frequency.....#
3.2	Revisions.....#
3.3	Archival Procedure.....#
4	Roles within the Scope of this Policy.....#
5	Related Legislation, References, and Documents.....#
6	Document Control.....#

2.1 Master and Reference Data Plan

1. The Entity shall develop, document, and maintain a Master and Reference Data Management Plan, ensuring alignment with the Entity's overall data strategy and governance objectives.
2. The MRDM Plan shall include a structured list of activities and milestones, incorporating both quick wins and long-term strategic initiatives to enhance the management and governance of Master and Reference Data.
3. All Master and Reference Data initiatives must have clearly defined owners and stakeholders, ensuring proper accountability and oversight for their implementation and ongoing maintenance.
4. The MRDM Plan shall identify and document interdependencies, assumptions, and risks to ensure a proactive approach in mitigating potential data management challenges.
5. A comprehensive implementation timeline shall be established within the MRDM Plan, detailing expected completion dates, budget allocations, and resource requirements to facilitate structured execution.
6. The MRDM Plan shall be monitored and periodically reviewed, with progress updates reported to the Data Governance Office (DGO) to ensure continuous alignment with evolving business and regulatory requirements.
7. Any significant changes or modifications to the MRDM Plan shall be formally reviewed and approved by the Data Governance Committee, with appropriate impact assessments conducted before implementation.

*** (Sample policy statements for a topic) ***

2.1.D
Cont.**Tip**

Develop the policy using clear, concise language, ensuring it addresses diverse scenarios while being tailored to the entity's needs. Ensure that each policy statement is directly connected to broader business objectives or outcomes, emphasizing how it contributes to achieving strategic goals and addressing the Entity's priorities. Focus on creating actionable, enforceable statements to drive consistent implementation and compliance.

2.1.E

Initial Review for Master and Reference Data Management Policy: The Data Management and Governance Section Head and the Data Strategy Expert review the preliminary policy to provide feedback and ensure compliance with the National Data Standards and other relevant regulations. This step ensures the policy aligns with external legal and regulatory requirements, safeguarding the Entity from potential non-compliance risks.

2.1.F

Incorporate Feedback: The Data Governance Specialist ensures the feedback has been implemented in the policy.

RACI Matrix – MRDM Policy Creation*Indicative*

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Metadata Specialist	Data Strategy Expert
2.1.A	Build Collaborative Forum	A	R	C	C	
2.1.B	Assess Current Master and Reference Data Management Policies		A	R	C	C
2.1.C	Determine Master and Reference Data Management Policy Updates		A	R	C	C
2.1.D	Develop/Update Master and Reference Data Management Policy		A	R	C	C
2.1.E	Initial Review for Draft Master and Reference Data Management Policy	A	R	I	C	R
2.1.F	Incorporate Feedback		A	R	C	

RACI Matrix Legend: R: Responsible A: Accountable C: Consulted I: Informed

Based on this Activity, the Entity can

Build Collaborative Forum of SMEs



Draft Master and Reference Data Management Policy

2.2 MRDM Policy Publication

This activity will guide the Entity on the steps to publish its data policies.

Process – MRDM Policy Publication

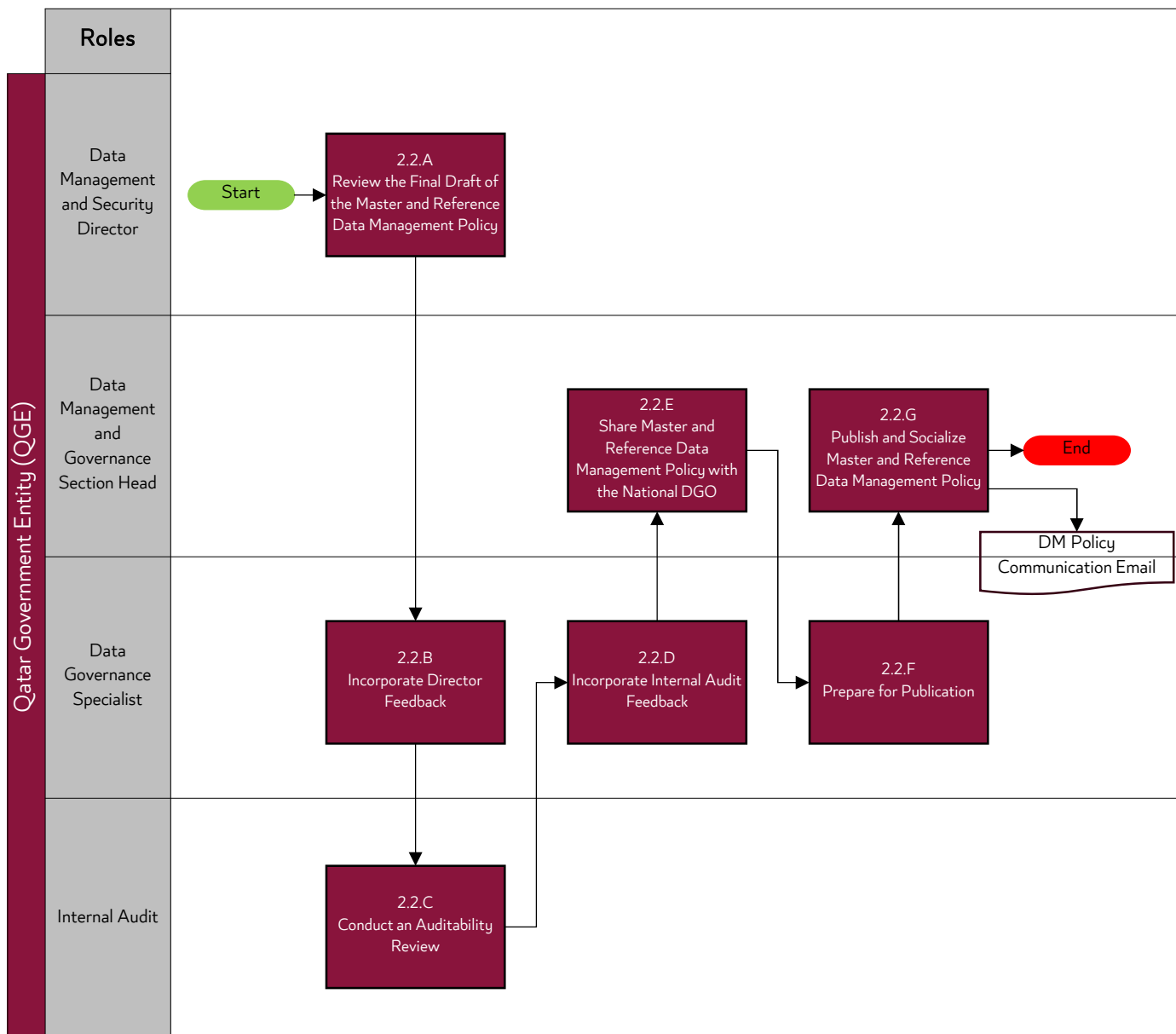


Figure 2: MRDM Policy Publication Process Map

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

2.2.A

Review the Final Draft of the Master and Reference Data Management Policy. The Data Management and Security Director reviews the final policy draft to ensure alignment with the entity's data management strategy and overarching business objectives. Any gaps or inconsistencies should be identified, with specific feedback provided for refinement.

2.2.B

Incorporate Director Feedback: Incorporate the feedback provided by the Data Management and Security Director, if any, to address any strategic misalignments or gaps identified during the review. This step ensures the policy reflects organizational priorities and strategic goals.

2.2.C

Conduct an Auditability Review: The draft policy is reviewed for auditability, ensuring it meets regulatory and operational standards. Feedback is gathered to address gaps in accountability, compliance, or performance tracking, if necessary.

2.2.D

Incorporate Internal Audit Feedback: Incorporate recommendations from the Internal Audit Team to ensure the policy is fully compliant with governance standards and operational integrity. This step is critical for maintaining transparency and accountability.

2.2.E

Share Master and Reference Data Management Policy with the National DGO: The Entity must submit the finalized draft of the Master and Reference Data Management Policy to the National DGO for their review and feedback. This step ensures that the policy aligns with national data program agenda and governance frameworks, regulatory requirements, and broader data management best practices.

2.2.F

Prepare for Publication: Incorporate all final feedback and prepare the policy for formal approval and publication. This includes ensuring consistency in language, structure, and alignment with both internal and external requirements.

2.2.G

Publish and Socialize Master and Reference Data Management Policy : Publish the finalized master and reference data management policy in a centralized, easily accessible repository, adhering to the Entity's established publication procedures.

Actively promote the policy across departments to ensure awareness and encourage adoption. Facilitate engagement sessions or workshops to familiarize teams with the policy's purpose and application. Consider how different roles and levels within the Entity are impacted by the policy change and tailor communications to clarify the expectations for each level. Collect feedback from stakeholders to assess its impact and identify any areas for further refinement or support in implementation.

**Tip**

Leverage surveys as an initial soft measure to assess engagement, awareness, and understanding of the published policy among key stakeholders. If engagement levels are low, strengthen communication efforts and gather further feedback to identify potential adjustments to the policy documentation before broader distribution.

 Asset Guidance

The **Data Management Policy Communication Email Template** is a sample announcement email the Entity can use to notify staff about the publication of the policy and provide key information.

Illustrative

Subject: [IMPORTANT] New [Domain Name] Policy Published – Please Review

Dear [Recipient Name/Team],

We are pleased to announce the publication of the new [Domain Name] Policy, which is now in effect as of [Date]. This policy is designed to [briefly explain the purpose, e.g., "ensure the effective management of metadata and data catalog operations across the entity"] and aligns with the National Data Standards.

What is the Policy About?

The [Domain Name] Policy governs the management, oversight, and compliance of [specific domain] activities across the entity. It outlines critical guidelines and procedures, including:

- [Key Point 1, e.g., "Data Management and Organizational Structure"]
- [Key Point 2, e.g., "Version control and security measures"]
- [Key Point 3, e.g., "Metadata quality and governance"]

This policy has been developed in collaboration with the Data Committee and relevant stakeholders, ensuring compliance with the national data management standards.

How Does This Affect You?

- **Action Required:** Please review the policy and ensure that your team is aligned with its requirements. The policy affects [mention departments or roles impacted, e.g., "all data stewards, IT staff, and department heads"]. Compliance with this policy is essential to ensure adherence to the National Data Management Standards, which are critical to the entity's data governance and operational success.
- **Changes to Processes:** [If applicable, briefly describe any changes to current workflows or responsibilities as outlined by the policy.]

Where Can You Access the Policy?

The policy is available on our internal [intranet/portal] and can be accessed through the following link:
[Link to Policy Document]

Next Steps

- **Review Deadline:** Please ensure that you and your team have reviewed the policy by [Date].
- **Training/Workshops:** [If applicable] We will be hosting a [training/workshop] on [Date] to assist teams in understanding and applying the new policy.
- **Ongoing Support:** For questions, feedback, or support, please contact [Support Contact] at [email/phone number].

Thank you for your attention to this important matter and for your cooperation in ensuring compliance with the new policy. Together, we can maintain a strong framework for [mention domain, e.g., "data catalog and metadata management"] across the entity.

Best regards,
[Your Name]
[Your Position]
[Your Department]
[Contact Information]

Process Risks: Develop and Publish MRDM Policy

Risk	Severity	Mitigation
1. Lack of engagement from business SMEs due to other demanding responsibilities.	High	Data Management and Governance Section Head to secure management commitment for the key SMEs required to develop the MRD policy.
2. No standardized or existing process to publish the policy.	Medium	Data Management and Governance Section Head to coordinate with the compliance office to establish a centralized process and repository for the publication and maintenance of MRD policy.

RACI Matrix – MRDM Policy Publication

Indicative

Sr No.	Activity	DGO	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Strategy Expert	Internal Audit
2.2.A	Review the Final Draft of Master and Reference Data Management Policy		A/R	C			
2.2.B	Incorporate Director Feedback		C	A	R		
2.2.C	Conduct an Auditability Review			I	I		A/R
2.2.D	Incorporate Internal Audit Feedback			A	R	C	
2.2.E	Share Master and Reference Data Management Policy with the National DGO	I	A	R			
2.2.F	Prepare for Publication			A	R	C	
2.2.G	Publish and Socialize Master and Reference Data Management Policy		A	R		C	

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Publish Master and Reference Data Management Policy

3. Master and Reference Data Management Planning



3. Master and Reference Data Management Planning

What is Master and Reference Data Management Planning and why does the Entity need it?

Master and Reference Data Management Planning is the structured process of defining, prioritizing, and implementing an Entity's approach to managing its critical data assets (e.g., customer, product, and supplier data) across various systems and departments. It aligns master and reference data management efforts with the Entity's overall data strategy, ensuring that key data attributes are prioritized, accurately maintained, and effectively governed. This planning process serves as the foundation for establishing a single source of truth across systems, supporting consistent data usage and reliability throughout the enterprise.

The need for MRDM Planning arises from several critical factors:

1. **Data Consistency:** Establishes a single source of truth and uniformity of data across departments, reducing errors from conflicting versions and ensuring all users have access to accurate data.
2. **Data Quality:** Provides processes to maintain high-quality data, addressing issues like duplicates and inaccurate records that can disrupt business operations.
3. **Compliance and Risk Management:** Helps meet regulatory requirements and reduces risks associated with incorrect or inconsistent data by establishing clear ownership and data controls.
4. **Operational Efficiency:** Centralizes data governance to reduce redundancies, streamline processes, and improve efficiency across departments.
5. **Data Integration and Interoperability:** Sets standards that enable seamless data integration across systems, supporting better data sharing and interoperability.

Implementing a structured MRDM plan allows Entities to make informed, data-driven decisions, improve operational efficiency, and ensure compliance with national data standards. It also strengthens organizational trust in data assets, making them more reliable and valuable in driving business performance and innovation.

The remaining sections of this chapter will detail the below processes that will guide the Entity on executing the below:



Guiding Principles for Master and Reference Data Management Planning

1. **Align with Business and Data Strategy:** Ensure the MRDM plan supports the Entity's strategic goals and delivers business value.
2. **Treat Data as an Asset:** Recognize master and reference data as valuable assets requiring governance and maintenance.
3. **Establish a Single Source of Truth:** Promote an authoritative source for master and reference data to reduce duplication and inconsistencies.
4. **Ensure Data Quality:** Maintain high standards for accuracy, completeness, and timeliness through data cleansing and validation.
5. **Define Ownership and Stewardship:** Assign clear roles for data custodianship to ensure accountability and alignment with governance.
6. **Balance Accessibility and Security:** Enable authorized access while protecting sensitive data with defined access rights and encryption.
7. **Standardize for Interoperability:** Create consistent definitions and formats for easy integration across systems.
8. **Plan for Scalability:** Design the MRDM plan to adapt to changing business needs, new data domains, and technologies.
9. **Promote Continuous Improvement:** Regularly review and enhance data management practices to keep the plan effective and relevant.
10. **Engage Stakeholders:** Involve key stakeholders and ensure clear communication for smooth adoption and alignment across departments.

3.1 MRD Discovery

MRD Discovery is a foundational step in the MRDM planning process. This activity will guide the entity in identifying, defining, and classifying MRD attributes (e.g., customer, products, supplier data) to ensure they are accurately sourced and effectively managed. By determining essential data domains and identifying authoritative sources, data owners, and data stewards, the entity establishes a single source of truth and clear accountability for maintaining data quality and accessibility.

3.1.A

Reference Existing Dataset and Attribute Inventory: If a dataset and attribute inventory has been established within the Data Catalog and Metadata Management domain, refer to it as the primary source for documenting master and reference data assets. Leverage this inventory to identify relevant datasets / attributes and avoid duplication of effort.

Asset Guidance

The **Data Catalog Template** includes a Dataset Inventory Template and an Attribute Inventory Template with pre-defined metadata to capture for the Entity's datasets and data attributes.

Illustrative

Dataset Inventory Template														
The following table reflects the Datasets' metadata required to describe datasets and tables														
Metadata Category	Business Metadata				Technical Metadata									
Metadata Attribute	Data Source ID	Data Source System	Dataset ID	Dataset Business Name (English)	Dataset Business Name (Arabic)	Dataset Business Description (English)	Dataset Business Description (Arabic)	Master Data Flag	Reference Data Flag	Business Tags	Update Frequency	Feedback Rating	Comment(s)	
Metadata Attribute ID	DS_BU_001	DS_BU_002	DS_BU_003	DS_BU_004	DS_BU_005	DS_BU_006	DS_BU_007	DS_BU_008	DS_BU_009	DS_BU_010	DS_BU_011	DS_BU_012	DS_BU_013	DS_BU_014
Requirement	Mandatory	Mandatory	Mandatory	Mandatory	Optional	Mandatory	Optional	Mandatory	Mandatory	Optional	Optional	Optional	Optional	Optional
Dataset Metadata														
Technical Metadata														
Source System Name	Table Name	Number of Rows	Qualified Name	Tracking Of Data Origin And Lineage	Documentation Of Data Modifications And The Entities Involved	Use Of Provenance Data In Compliance And Auditing	Technical Use Of Provenance Data In Compliance And Auditing	Disaster Recovery Time Objective (RTO)	Disaster Recovery Point Objective (RPO)	Disaster Recovery Plan Testing Schedule	Backup Frequency	Backup Retention Period	Backup Location	
DS_TC_001	DS_TC_002	DS_TC_003	DS_TC_004	DS_TC_005	DS_TC_006	DS_TC_007	DS_TC_008	DS_TC_009	DS_TC_010	DS_TC_011	DS_TC_012	DS_TC_013	DS_TC_014	DS_TC_015
Optional	Mandatory	Optional	Mandatory	Optional	Optional	Optional	Optional	Optional	Optional	Optional	Optional	Optional	Optional	Optional
Data Ownership Metadata														
Authoritative Data Source	Data Owner Full Name	Data Owner Email Address	Data Owner Phone Number	Data Owner Department	Business Data Steward Full Name	Business Data Steward Email	Business Data Steward Phone Number	Business Data Steward Department	It Data Steward Full Name	It Data Steward Email Address	It Data Steward Phone Number	It Data Steward Department		
DS_DO_001	DS_DO_002	DS_DO_003	DS_DO_004	DS_DO_005	DS_DO_006	DS_DO_007	DS_DO_008	DS_DO_009	DS_DO_010	DS_DO_011	DS_DO_012	DS_DO_013	DS_DO_014	DS_DO_015
Mandatory	Mandatory	Mandatory	Optional	Optional	Mandatory	Mandatory	Optional	Optional	Mandatory	Mandatory	Optional	Optional	Optional	Optional
Data Quality Metadata														
Accuracy	Completeness	Consistency	Timeliness	Validity	Uniqueness	Data Classification	Data Retention Period	Data Anonymization Techniques	Data Encryption Status	Data Masking	Access Control Level	Compliance Requirements	Breach Notification Contact	
DS_DQ_001	DS_DQ_002	DS_DQ_003	DS_DQ_004	DS_DQ_005	DS_DQ_006	DS_DQ_007	DS_DQ_008	DS_DQ_009	DS_DQ_010	DS_DQ_011	DS_DQ_012	DS_DQ_013	DS_DQ_014	DS_DQ_015
Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Optional	Optional	Optional	Optional	Optional	Optional	Optional	Optional

Attributes Inventory Template														
The following table reflects the Attributes' metadata required to describe attribute-level data														
Metadata Category	Business Metadata				Technical Metadata									
Metadata Attribute	Data Source ID	Data Source System	Dataset ID	Dataset Business Name (English)	Dataset Business Name (Arabic)	Dataset Business Description (English)	Dataset Business Description (Arabic)	Critical Data Element	List Of Values	Master Data Flag	Reference Data Flag	Reference ID	Reference Data Name	Business Tags
Metadata Attribute ID	ATT_TC_001	ATT_TC_002	ATT_TC_003	ATT_TC_004	ATT_TC_005	ATT_TC_006	ATT_TC_007	ATT_TC_008	ATT_TC_009	ATT_TC_010	ATT_TC_011	ATT_TC_012	ATT_TC_013	ATT_TC_014
Requirement	Mandatory	Mandatory	Mandatory	Mandatory	Optional	Mandatory	Optional	Mandatory	Optional	Mandatory	Optional	Mandatory	Optional	Optional
Attribute Metadata														
Technical Metadata														
Attribute ID	Table Name	Column Name (English)	Column Description (English)	Technical Data Type	Data Length	Data Format	Character Set	Source Of List Of Values Standard	PII/FIN/	Nullability (Y/N)	Version Number	Version Updated By	Version Update Date	Authoritative Data Source
ATT_TC_001	ATT_TC_002	ATT_TC_003	ATT_TC_004	ATT_TC_005	ATT_TC_006	ATT_TC_007	ATT_TC_008	ATT_TC_009	ATT_TC_010	ATT_TC_011	ATT_TC_012	ATT_TC_013	ATT_TC_014	ATT_TC_015
Mandatory	Mandatory	Optional	Optional	Mandatory	Optional	Mandatory	Mandatory	Optional	Optional	Mandatory	Mandatory	Optional	Mandatory	Mandatory
Data Privacy & Security Metadata														
Data Classification	Data Anonymization Techniques	Data Encryption Status	Data Masking	Access Control Level	Compliance Requirements	Breach Notification Contact								
ATT_DP_001	ATT_DP_002	ATT_DP_003	ATT_DP_004	ATT_DP_005	ATT_DP_006	ATT_DP_007								
Mandatory	Optional	Optional	Optional	Optional	Optional	Optional								

3.1.B

Create Data Inventory (If None Exists): If a dataset or attribute inventory is not readily available within the Data Catalog, refer to the Data Catalog and Metadata Management QDKC to create the necessary inventory. This step ensures the inventory accurately captures all required master and reference data, establishing a strong foundation for master and reference data efforts.

3.1.C

Define Data Domain Categories: The Data Governance Specialist collaborates with the respective Data Stewards to establish specific data domain categories for master and reference data based on the Entity's needs and data management objectives. Common categories may include customer, product, supplier, location, and financial data. Define each category to ensure consistent understanding across the Entity.

3.1.D

Tag Datasets Against Defined Master and Reference Data Domains: Identify and tag each dataset according to its relevant data domain (e.g., customer, product). This categorization ensures that datasets are clearly aligned with their functional domain, simplifying governance and management practices.

3.1.E

Classify Datasets and Attributes as Master or Reference Data: Within the Data Catalog, classify each dataset and attribute in the respective inventory as either master or reference data based on its role in supporting core business processes.

3.1.F

Define Authoritative Source for Each Dataset: Identify the authoritative source for each dataset, particularly where multiple sources of similar data exist. Designating an authoritative source for each dataset within the inventory minimizes inconsistencies and establishes a single, trusted source of truth for data domains.

3.1.G

Categorize Datasets as Internal or External Data: Categorize each dataset based on its origin:

- Internal (originating within the entity)
- External (sourced from an external entity)

This distinction supports the data governance and data sharing, integration, and interoperability domains by highlighting dependencies on external sources.

3.1.H

Outline Dataset Owners and Stewards: Assign a data owner and a data steward for each dataset in the inventory. The data owner is accountable for policies, access, and alignment with business goals, while the data steward manages day-to-day data quality, consistency, and adherence to governance standards. Clearly designating these roles ensures accountability and supports high data quality and effective data use across the Entity.

3.1.I

Validate Inventory with Data Owners: Present the master and reference data inventory to data owners and stakeholders from each domain. Obtain feedback to confirm that the inventory is accurate, comprehensive, and aligned with business needs. This validation process ensures that the inventory reflects the priorities and requirements of those responsible for data governance.

3.1.J

Approve or Revise Inventory Based on Feedback: Based on feedback from data owners, either approve the inventory or revise it, as necessary. If revisions are required, implement feedback to address any gaps or inaccuracies, and repeat the validation process as needed to ensure the inventory is complete and accurate.



Key Message / Information

Maintaining an up-to-date Data Catalog as the central repository for dataset inventories is essential for consistent management of master and reference data. The catalog should define master and reference data domains, authoritative sources, and custodianship roles in line with National Data Management Standards. Each dataset must be systematically tagged, categorized, and validated with input from data owners and stewards, establishing a trusted "single source of truth" across the Entity. This approach enhances data quality, governance, interoperability, and compliance.

RACI Matrix – MRD Discovery

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Metadata Specialist	Data Owner	Data Steward
3.1.A	Reference Existing Dataset and Attribute Inventory		A	R	C		
3.1.B	Create Data Inventory (If None Exists)		A	R	C		
3.1.C	Define Data Domain Categories	I	A	R			C
3.1.D	Tag Datasets Against Defined Master and Reference Data Domains		A	R			C
3.1.E	Classify Datasets and Attributes as Master or Reference Data		A	R			C
3.1.F	Define Authoritative Source for Each Dataset	I	A	R			
3.1.G	Categorize Datasets as Internal or External Data		A	R			C
3.1.H	Outline Dataset Owners and Stewards	I	A	R		C/I	C/I
3.1.I	Validate Inventory with Data Owners		A	R		C	
3.1.J	Approve or Revise Inventory Based on Feedback	I	A	R		I	I

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Define Data Domain Categories



Define Authoritative Source for its datasets



Outline Dataset Owners and Stewards

3.2 MRDM Requirements Gathering

This section outlines the process for gathering and prioritizing MRDM requirements to ensure alignment with Entity objectives. It includes identifying key data needs, consolidating stakeholder inputs, and prioritizing Critical Data Elements (CDEs) based on their importance to decision-making, operational efficiency, and regulatory compliance.

By capturing and prioritizing these requirements, entities can establish a strong foundation for designing and implementing an MRDM solution that addresses both functional and technical needs while supporting long-term governance and data quality goals.

3.2.1 Capture MRDM Requirements

This section outlines the steps to capture MRDM requirements for effectively managing the Entity's MRD across the data lifecycle.

Process – Capture MRDM Requirements

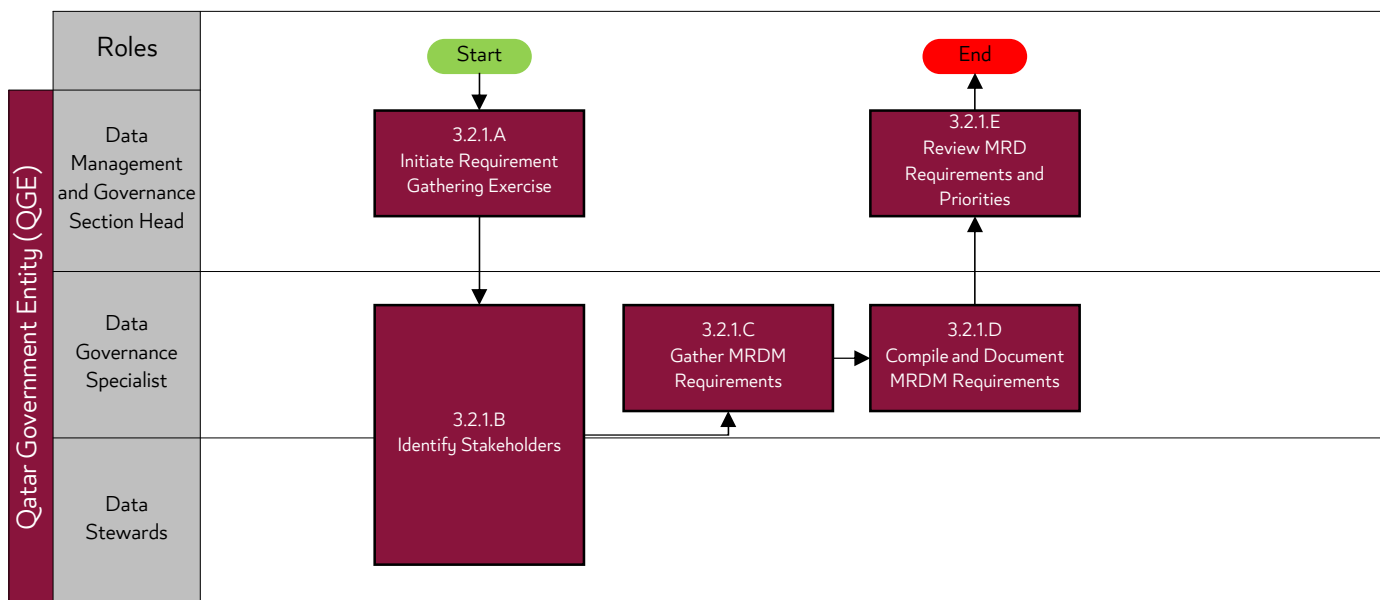


Figure 3: Process Map for Capturing MRDM Requirements

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

3.2.1.A

Initiate Requirement Gathering Exercise: The Data Management and Governance Section Head initiates the MRDM requirement gathering exercise by defining its scope, objectives, and assigning the Data Governance Specialist to coordinate the activity. A high-level plan should be developed to guide the requirements collection process effectively and align with the Entity's MRDM goals.



Key Message / Information

Clearly defined objectives reduces ambiguities and ensures all stakeholders are aligned on the purpose and outcomes of the requirement-gathering process.

3.2.1.B

Identify Stakeholders: The Data Governance Specialist identifies key stakeholders essential for the MRDM requirement gathering exercise. Stakeholders include, but are not limited to:

- **Data Owners:** Ensure ownership, accountability, and governance of MRD, aligning it with business needs and operational goals.
- **Data Stewards (Business and IT):** Represent functional and technical needs for managing and utilizing Master & Reference Data.
- **System Owners (Source Systems and Downstream Systems):** Responsible for systems where Master & Reference Data (MRD) is created, read, updated, deleted, and consumed.
- **Data Compliance Specialist:** Ensure MRD aligns with security policies, privacy regulations, and compliance frameworks.
- **Data Strategy Expert:** Align MRDM practices with the Entity's data strategy and broader business objectives.
- **Business Data Analysts:** Provide insights into data consumption patterns and reporting needs to enhance decision-making.
- **Data Architects:** Define and maintain the overarching data architecture and data models for MRD, including the design and integration of the MDM Hub.

3.2.1.C
(1/3)

Gather MRDM Requirements: The Data Governance Specialist facilitates workshops with identified stakeholders to collect MRDM requirements. These discussions ensure that stakeholder needs are fully captured and provide a collaborative platform to uncover challenges, identify opportunities, and document requirements for managing MRD.

Workshops and meetings should focus on understanding data creation, usage, governance processes, and address any potential gaps in current MRDM practices. Below are key activities to consider when preparing and running the meetings and workshops:

i. **Meeting and Workshop Preparation:**

- Develop a detailed agenda, including topics of discussion tailored to each group of stakeholders (e.g., Data Owners, Data Stewards, System Owners etc.).
- Share the agenda and preparatory material in advance to ensure participants are well-informed and ready to contribute.

ii. **Tailor Questions to Stakeholders:**

- **Data Owners:** Focus on ensuring clear ownership and accountability for MRD, aligning it with business needs and operational goals. Oversee data availability, accessibility, and governance to maintain MRD as a single source of truth. Define policies for data validation, lifecycle management, and cross-entity collaboration.
- **Data Stewards (Business and IT):** Focus on ensuring the consistency, accuracy, and governance of MRD across systems. Discuss challenges related to data validation, quality monitoring, and lifecycle management, ensuring data remains reliable and aligned with governance policies. Explore best practices for maintaining MRD integrity across business functions, enforcing data stewardship workflows, and addressing cross-system data consistency issues.

3.2.1.C
Cont.
(2/3)

- **System Owners (Source Systems and Downstream Systems):** Focus on technical requirements for systems where MRD is created, stored, and consumed. Explore integration challenges, performance requirements (e.g., scalability), and data flow inefficiencies.
 - **Data Compliance Specialists:** Focus on compliance frameworks, privacy regulations, and security concerns. Explore audit and reporting capabilities, as well as requirements for data masking, encryption, and adherence to security standards.
 - **Data Strategy Experts:** Focus on aligning MRDM initiatives with the broader data strategy and strategic business drivers. Explore how MRDM can support cross-departmental collaboration and enable desired outcomes like improved citizen services or operational efficiency.
 - **Business Data Analysts:** Focus on leveraging MRD for data-driven decision-making, ensuring structured data supports analytics, reporting, and business intelligence. Identify gaps, enhance usability, and advocate for governance practices that improve data reliability.
 - **Data Architects:** Focus on the data architecture standards and the common architecture archetypes (e.g., centralized, coexistence) for MRD. Explore metadata, data lineage, and governance requirements, as well as handling legacy systems or silos.
- iii. **Capture Input on Key Topics:** Collect input on the following example areas to build a comprehensive understanding of MRDM requirements:
- **Strategic Requirements:** Identify how MRDM initiatives align with the Entity's strategic goals and objectives. Discuss business drivers such as enhancing citizen services, operational efficiency, or compliance with regulatory mandates. Capture metrics or outcomes stakeholders expect to achieve. At a minimum, input on these topics should be gathered from Data Owners and Data Strategy Experts.
 - **Governance Requirements:** Determine the processes and procedures needed to manage, oversee, and govern MRD (e.g., data creation, modification, and archiving). At a minimum, insights can be provided by the Data Governance Specialist and Data Quality Specialists.
 - **Data Standards:** Discuss the high-level guidelines required to ensure best practices in MRDM. High-level guidelines include but are not limited to defining naming conventions, data formats, metadata structures. At a minimum, contributions should come from Data Stewards, Data Architects, and Data Compliance Specialists.
 - **Hub Architecture:** Define requirements for the MDM Hub's design, structure, and operations. Discuss preferred architecture models (e.g., centralized, coexistence, registry) and how data storage, creation, updates, and distribution should be managed. At a minimum, input should be gathered from Data Stewards, Data Architects and System Owners.
 - **Integration & Interoperability:** Explore the connection and data exchange requirements between the MDM Hub and other systems. Capture integration methods (e.g., APIs, batch processing), required data flows, and challenges with existing system dependencies. At a minimum, these discussions should involve Data Stewards, Data Architects and System Owners.

3.2.1.C
Cont.
(3/3)

- **Performance Requirements:** Identify expectations for system performance, including processing speed, scalability, and responsiveness. Discuss anticipated data volumes, system uptime requirements, and capacity for future growth. Input on these topics should be sought from System Owners, Operations and Maintenance Teams, and Data Architects.
- **Data Provisioning:** Gather requirements for delivering MRD to consuming systems to support operational and analytical needs. For example, provisioning data might involve scheduling daily exports of master data to a Data Warehouse (DWH) for reporting purposes. Relevant insights can be provided by Data Stewards, Data Architects and System Owners.
- **Security & Privacy:** Define requirements to ensure MRD is protected through robust access controls, encryption, and privacy measures. Discuss compliance with legal, adhering to regulatory standards, and address specific concerns around sensitive data handling. Contributions should come from Data Compliance Specialists, System Owners, and Data Architects.
- **Metadata & Lineage:** Capture requirements for tracking metadata and data lineage. Discuss how metadata should be structured, maintained, and updated to ensure governance, transparency, and trust in MRD. Input on these topics should be gathered from Data Stewards, Data Architects, and Data Quality Specialists.
- **Analytics & Reporting:** Identify reporting and analysis needs for monitoring data usage, performance, and trends. Capture specific metrics or insights stakeholders require to support decision-making and assess the effectiveness of MRDM practices. Relevant insights can be provided by Business Data Analysts, and Data Strategy Experts.

3.2.1.D

Compile and Document MRDM Requirements: After identifying stakeholders and gathering requirements across the various MRDM requirement types, the next step is to compile and document these requirements systematically. This ensures that all collected input is organized and recorded for use in planning, prioritization, and implementation. Entities should use the MRDM Requirements Template as the standard tool for documenting these requirements. The steps to compile and document MRDM requirements are as follows:

i. **Review Collected Input:**

- Consolidate all the information gathered from workshops, meetings, and stakeholder discussions.
- Ensure that each requirement is categorized accurately by its type (e.g., Strategic, Governance, Data Standards).
- Cross-check inputs to identify duplicates or inconsistencies that need clarification.

ii. **Populate the MRDM Requirements Template:**

- Use the provided template to document each requirement systematically.
- Fill out the following fields for each requirement:
 - **Requirement ID:** Assign a unique identifier to each requirement.
 - **Requirement Type:** Select the appropriate type from the predefined list.
 - **Challenge:** Provide a concise description of the challenge the requirement addresses (if applicable).
 - **Requirement:** Document a clear, actionable statement of the requirement.
 - **Priority:** Categorize the consolidated requirements into High, Medium, or Low priority. This prioritization helps identify critical requirements to address in subsequent architectural and implementation phases.



Key Message / Information

Prioritization ensures the Entity focuses resources on high-impact requirements while planning for future enhancements.

- **Requirement Requested By:** Record the name and role of the individual or team that identified the requirement.
- **Requirement Owner:** Record the name and role of the individual or team responsible for overseeing the requirement's implementation.
- **Dependencies:** Document dependencies that must be considered to fulfill the requirement.
- **Date Collected:** Document the date that the requirement was collected.

iii. **Ensure Completeness and Clarity:**

- Verify that all fields in the template are complete for each requirement.
- Ensure that requirements are written in clear, unambiguous language to avoid misinterpretation.
- Seek clarification from stakeholders if any information is incomplete or unclear.

3.2.1.D
Cont.

iv. Validate Documented Requirements:

- Share the completed MRDM Requirements Template with key stakeholders, such as the Data Management and Governance Section Head and relevant team leads, for review and validation.
- Incorporate feedback to refine and finalize the documented requirements.

 Asset Guidance

The MRDM Requirements Template includes an MRDM Requirements Sheet, which the Entity can use to detail specific master and reference data management requirements.

Illustrative

MRDM Requirements	
The following table can be used to document the overall master and reference data management challenges and requirements within the Entity	

Field	Description
Requirement ID	Unique number to distinctly identify the requirement.
Requirement Type	Specify (via the drop down list) what type of requirement is being documented. 1. Strategic: High-level requirements that define the alignment of MDM with organizational goals and objectives. 2. Governance: Requirements focused on defining the policies, roles, and processes to manage, oversee, and govern MRD. 3. Data Standards: Requirements addressing the data standards needed to standardize master and reference data. 4. Hub Architecture: Requirements defining the design, structure, and operation of the MDM Hub, including where and how MRD will be stored, created, or updated. 5. Integration & Interoperability: Requirements related to the connection and data exchange between the MDM Hub and other systems, including what data needs to be integrated and how it flows. 6. Performance Requirements: Requirements outlining the processing speed, scalability, and responsiveness needed to ensure the MDM Hub meets organizational demands. 7. Data Provisioning: Requirements specifying what data needs to be made available, to which consuming systems, and how it will be delivered (e.g., APIs, batch jobs). 8. Security & Privacy: Requirements defining how master and reference data will be protected through access control, encryption, and compliance measures. 9. Metadata & Lineage: Requirements specifying how metadata will be captured, maintained, and how data lineage will be tracked for governance and transparency. 10. Analytics & Reporting: Requirements related to monitoring, reporting, and analyzing data usage and trends.
Challenge	Description of the challenge in managing master or reference data effectively. E.g., duplicate citizen records across systems.
Requirement	A detailed statement outlining the requirement to achieve a specific objective.
Priority	The urgency or importance of addressing the requirement or challenge.
Requirement Requested By (Name)	The name of the individual or team that identified the requirement or challenge.
Requirement Requested By (Role)	The role of the individual or team that requested the requirement (e.g., Data Steward, Business Analyst).
Requirement Owner (Name)	The name of the individual or team responsible for implementing or overseeing the requirement.
Requirement Owner (Role)	The role of the individual or team managing the requirement's implementation (e.g., IT Lead, Data Governance Lead).
Dependencies	Dependencies that must be taken into account to fulfill the requirement.
Data Collected	Date on which the requirement was collected.
Comments	Any special considerations or notes related to the requirement.

Requirement ID	Requirement Type	Challenge (Optional)	Requirement
MRDM.R1	Strategic	E.g., Currently no single source of truth for master data.	E.g., Establish a single source of truth for master data
MRDM.R2	Governance	E.g., No defined approval workflow for creating or modifying master data.	E.g., Define approval workflows for creating or modifying master data
MRDM.R3	Data Standards	E.g., Customer data is not standardized.	E.g., Develop standardization guidelines for customer master data
MRDM.R4	Hub Architecture	E.g., NA	E.g., Two-way synchronization between MDM Hub and source systems
MRDM.R5	Integration & Interoperability	E.g., NA	E.g., Real-time API integration with source systems (e.g., ERP, CRM).
MRDM.R6	Performance Requirements	E.g., NA	E.g., Ensure real-time synchronization latency does not exceed 5 seconds.
MRDM.R7	Data Provisioning	E.g., NA	E.g., Schedule daily exports of master data to a DWH for reporting purposes.
MRDM.R8	Security & Privacy	E.g., NA	E.g., Restrict access to sensitive data attributes (e.g., salary information, national ID).
MRDM.R9	Metadata & Lineage	E.g., No traceability of modifications made to master data.	E.g., Maintain a lineage trail showing how master data attributes have been derived or modified.
MRDM.R10	Analytics & Reporting	E.g., NA	E.g., Provide insights into data update trends for operational monitoring.

Priority	Requirement Requested By (Name)	Requirement Requested By (Role)	Requirement Owner (Name)	Requirement Owner (Role)	Dependencies	Date Collected	Comments

3.2.1.E

Review MRD Requirements and Priorities: The Data Management and Governance Section Head reviews the consolidated and prioritized MRD requirements to ensure alignment with the Entity's data management strategy, regulatory obligations, and business objectives.

RACI Matrix – Capture MRDM Requirements

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Stewards	Data Owners
3.2.1.A	Initiate Requirement Gathering Exercise	A	R	I		
3.2.1.B	Identify Stakeholders		A	R	R	
3.2.1.C	Conduct Meetings and Workshops to Collect Business Requirements		A	R	C	C
3.2.1.D	Compile and Document MRDM Requirements		A/C	R	C	C
3.2.1.E	Review MRD Requirements and Priorities	A	R			

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Identify MRDM Requirements

3.2.2 Capture Attribute-Level MRD Requirements

This section guides the Entity in capturing attribute-specific requirements that directly support the identification and prioritization of CDEs. These requirements serve as the foundation for developing attribute-level trust, similarity, and survivorship rules.

Process – Capture Attribute-Level MRD Requirements

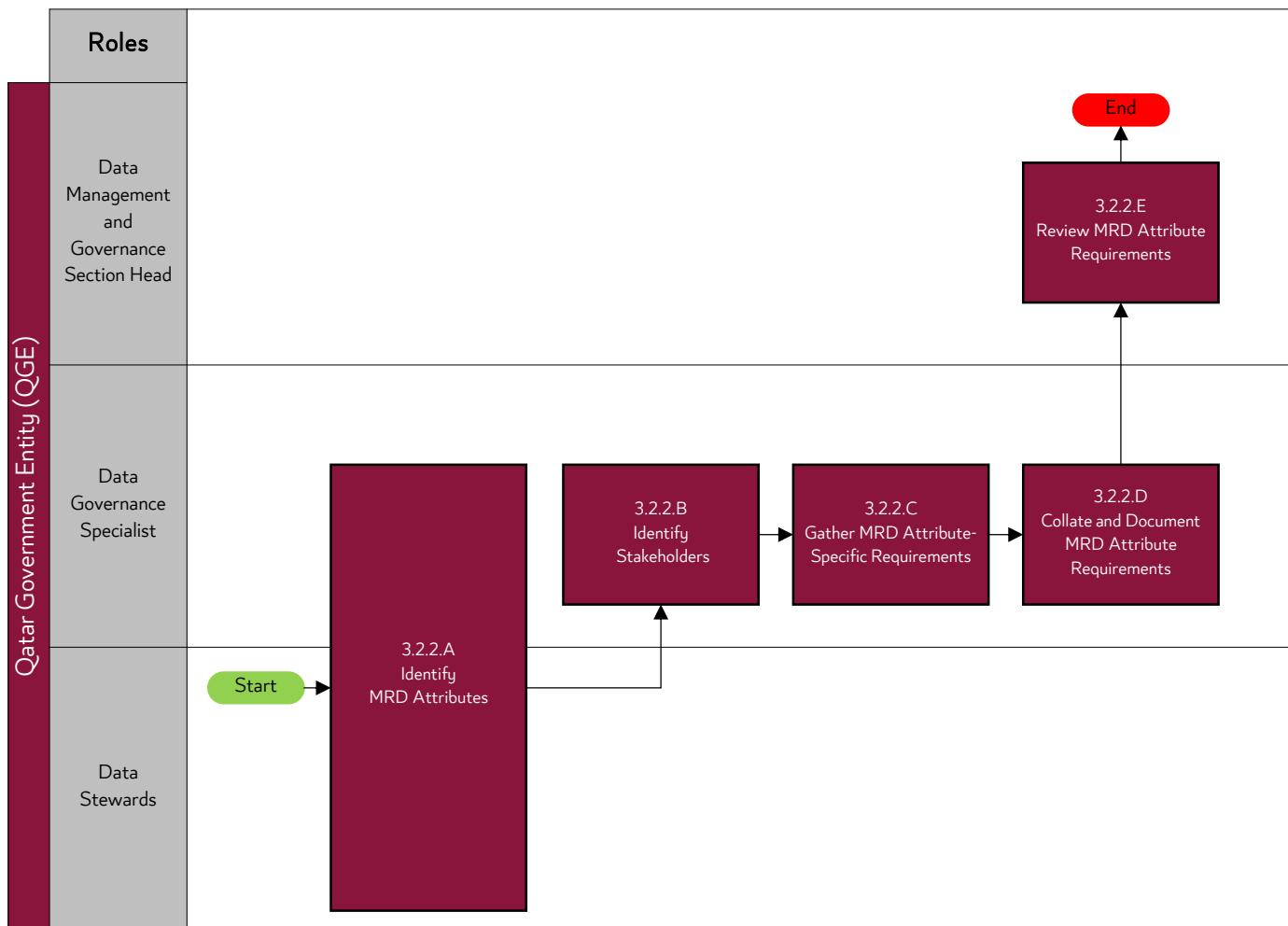


Figure 4: Process Map for Capturing Attribute-Level MRD Requirements

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

3.2.2.A

Identify MRD Attributes: The Data Governance Specialist works with Data Stewards to identify MRD attributes from the documented inventory within the Data Catalog. This includes identifying key datasets (e.g., Customer, Employee, Product) and their associated attributes (e.g., Name, Address, Citizen ID).



Key Message / Information

Identifying key attributes ensure the focus remains on the most critical data assets for the Entity's operations and decision making.

3.2.2.B

Identify Stakeholders: The Data Governance Specialist identifies key stakeholders essential for the MRD attribute-level requirement gathering exercise by leveraging the data catalog to support in the identification of stakeholders who manage, use, or depend on the MRD. Key stakeholders include but are not limited to:

- **Data Owners and Business Stakeholders:** For business and operational needs.
- **Data Stewards:** For data governance and quality concerns.
- **System Owners:** For technical input related to system dependencies.

**Key Message / Information**

Pre-identifying attributes and stakeholders ensures that workshops are focused, actionable, and involve the right participants.

3.2.2.C

Gather MRD Attribute Requirements: With the preliminary list of attributes and stakeholders defined in the previous steps, conduct workshops to gather detailed attribute-level requirements. Focus on operational, technical, and governance aspects of each attribute, including validation rules, dependencies, and quality expectations. Below are key activities to consider when preparing and running the workshops:

i. Workshop Preparation:

- Develop a detailed agenda, including topics of discussion and questions tailored to each group of stakeholders (e.g., System Owners, Data Stewards).
 - **Data Owners and Business Stakeholders:** Focus on business challenges and data-related pain points (e.g., inconsistencies in citizen records across systems, inaccurate reports).
 - **Data Stewards:** Address operational challenges, such as data quality issues, frequency of updates, and high-level rules for data matching and merging.
 - **System Owners:** Focus on technical requirements (e.g., data integration, system dependencies, and governance needs).
- Share the agenda and preparatory material in advance to ensure participants are well-informed and ready to contribute.

ii. Workshop Execution:

- Gather attribute-specific requirements. Key considerations include:
 - **High-level Validation Methods and Rules:** Formats, standardization, and range checks (e.g., Phone Number must comply with the E.164 standard).
 - **Dependencies:** Which systems create or consume the data (e.g., CRM as the source for Customer Name).
 - **Usage and Frequency:** How often the data is updated, accessed, or shared.

3.2.2.D

Collate and Document MRD Attribute Requirements: Populate the **MRD Attribute Requirements Template** with the attribute-specific requirements captured during the workshops.

Asset Guidance

The **MRDM Requirements Template** includes an **MRD Attribute Requirement Sheet**, which the Entity can use to detail MRD Attribute Requirements.

Illustrative

MRD Attribute Requirements	
The following table can be used to document the MRD requirements in detail	

Field	Description
Requirement ID	Unique number to distinctly identify the requirement.
Master Data Domain	The master data domain identified for the Entity (e.g., Customer, Vendor, Product).
Data Owner	Name of the individual responsible for the master/reference data.
Data Steward	Name of the data steward responsible for maintaining the data.
Attribute Name	The master data identified within the domain or the reference data linked to it (e.g., Address).
MRD Classification	The classification of the attribute (master or reference data).
Attribute Business Challenge	The challenge related to the master or reference data attribute.
Attribute Usage Details	Details of how the master or reference data will be used (e.g., supporting compliance reporting, enabling analytics, streamlining operational processes).
Attribute Requirement Details	Capture the specific attribute requirements, covering the following sample areas: 1. Data Acquisition: How the data will be acquired from the source system (e.g., delta load, initial load). 2. Data Quality Requirements: High-level rules for match/merge, data quality checks, etc. 3. Data Distribution: Distribution methods (e.g., API, batch jobs). 4. User Role Management: Requirements for role-based access and authentication (e.g., LDAP). 5. Other Requirements.
Consuming Systems	Downstream systems consuming the master/reference data (e.g., BI tools, transactional systems).
Dependency	Dependencies on other master or reference data, if any.
Requirement Provided By	Name, department, and role of the individual who provided the input.
Data Collected	Date on which the requirement was collected.
Comments	Any special considerations or notes related to the requirement.

Requirement ID	Master Data Domain	Data Owner	Data Steward	Attribute Name	MRD Classification	Attribute Business Challenge
MRDR1	Customer	Ahmed Sakr	Mohamed Majid	Residential Address	Master Data	Ensuring accurate and up-to-date email address records for customer communication and service delivery.

Attribute Usage Details	Attribute Requirements Details	Consuming Systems	Dependency	Requirement Provided By	Date Collected	Comments
Used to contact customers for service delivery, validate billing addresses, and generate location-based analytics.	Deduplicate and validate email address for correct syntax.	ERP, Logistics System, Data Warehouse	1. Reference Data - Contact Detail Type 2. Customer master	Abdul Ahmed	05.01.2025	



Key Message / Information

Once the MRD attributes have been identified, leverage the Data Catalog to populate the Master Data Domain, Attribute Name, MRD Classification, Data Owner, Data Steward, and Data Source columns within the MRD Attribute Requirements Template.

When specifying MRD requirements, ensure that all stakeholder inputs align with the National Data Standards. For example, ensure that any data quality requirements comply with the National Data Quality Management Standards. This approach promotes consistency, compliance, and alignment across all aspects of MRDM.

3.2.2.E

Review MRD Attribute Requirements: The Data Management and Governance Section Head reviews the consolidated MRD attribute requirements to ensure alignment with the Entity's data management strategy, business objectives, and regulatory obligations.

RACI Matrix – Capture Attribute-Level MRD Requirements

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Stewards	Data Owners
3.2.2.A	Identify MRD Attributes		A	R	R	I
3.2.2.B	Identify Stakeholders	I	A	R	C	I
3.2.2.C	Gather MRD Attribute-Specific Requirements		A	R	C	I
3.2.2.D	Collate and Document MRD Attribute Requirements	I	A	R	C	I
3.2.2.E	Review MRD Attribute Requirements	A	R			

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Identify Attribute-Specific MRD Requirements

3.2.3 Identify Prioritized Attributes (CDEs)

Identifying CDEs is essential for focusing MRDM on the most impactful assets within the Entity. CDEs are key attributes that significantly influence decision-making, operational efficiency, regulatory compliance, and risk management. This activity guides the Entity on steps to identify and prioritize CDEs from the documented MRD inventory within the Data Catalog.

3.2.3.A

Identify CDEs: Review the documented MRD inventory within the Data Catalog in addition to the MRD attribute requirements to identify CDEs defined during the data strategy, cataloging, and requirements gathering exercises. If required, address potential gaps by identifying additional CDEs based on their usage, process importance, or regulatory compliance needs, focusing on elements critical to decision-making, operational efficiency, compliance, and risk management. Update the Data Catalog and MRD Attribute Requirements Template as needed.



Key Message / Information

Identifying key attributes ensure the focus remains on the most critical data assets for the Entity's operations and decision making.

3.2.3.B

Evaluate CDEs Using the CDE Prioritization Framework: Evaluate each identified CDE using the criteria in the CDE Prioritization Framework asset. Assess each attribute on factors including Business Criticality, Usage Frequency, Data Quality Issues, Integration, and Change Frequency. Assign a score from 1 (low) to 5 (high) across all criteria, and document data source linkage, table linkage, and business metadata for traceability and context.



Asset Guidance

The **CDE Prioritization Framework** provides a weighted criteria to help the Entity prioritize CDEs. Based on the assigned weighted score for each criteria, the template automatically calculates a total score.

Illustrative



المجلس الوطني للتخطيط
National Planning Council
دولة قطر - State of Qatar

CDE Prioritization Framework

The following table reflects the criteria and respective weights to prioritize Attributes (CDEs)

Criteria	Weight (%)	Description	Score (Range 1-5)
Business Criticality	30%	How critical is this attribute for business operations, the data strategy, decision-making, or public services?	1 = Low importance, 5 = High importance
Usage Frequency	20%	How often is this attribute used across systems and processes?	1 = Rarely used, 5 = Frequently used
Data Quality Issues	25%	Is this attribute prone to data quality issues such as inconsistency, inaccuracy, or incompleteness?	1 = No issues, 5 = Significant issues
Integration and Data Sharing	20%	How important is this attribute for data integration and sharing between different systems/entities?	1 = Low importance, 5 = High importance
Change Frequency (Volatility)	5%	How often does this attribute change (e.g., dynamic or static data)?	1 = Rarely changes, 5 = Frequently changes

Prioritization						
Business Criticality (30%)	Usage Frequency (20%)	Data Quality (25%)	Integration (20%)	Volatility (5%)	Total Score	Final Priority
					0	low
					0	low
					0	low

3.2.3.C

Validate CDE List with Stakeholders: Engage data owners, stewards, and stakeholders in reviewing the prioritized CDE list to ensure that it aligns with strategic objectives, operational needs, and regulatory compliance and risk management requirements.

RACI Matrix – Identify Prioritized Attributes (CDEs)

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Owner	Data Steward
3.2.3.A	Identify CDEs	I	A	R		C
3.2.3.B	Evaluate CDEs Using the CDE Prioritization Framework	I	A	R		C
3.2.3.C	Validate CDE List with Stakeholders	I	A	R	C	C

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Identify and Evaluate CDEs

3.3 MRDM Plan

Based on the master and reference data discovered and prioritized in the previous activity, a master and reference data management plan is to be developed and aligned with the Entity's data strategy to guide the implementation of master and reference data management initiatives across the Entity. The plan should outline key activities and milestones, addressing both quick wins and longer-term strategic objectives. Additionally, the plan will designate initiative owners and key stakeholders, ensuring accountability and collaborative support throughout the implementation process. It will also highlight interdependencies with other projects, assumptions that underpin the plan, and potential risks to be mitigated. A detailed implementation timeline will be developed, along with a structured cost breakdown to support resource allocation and budgeting.

Process – MRDM Plan

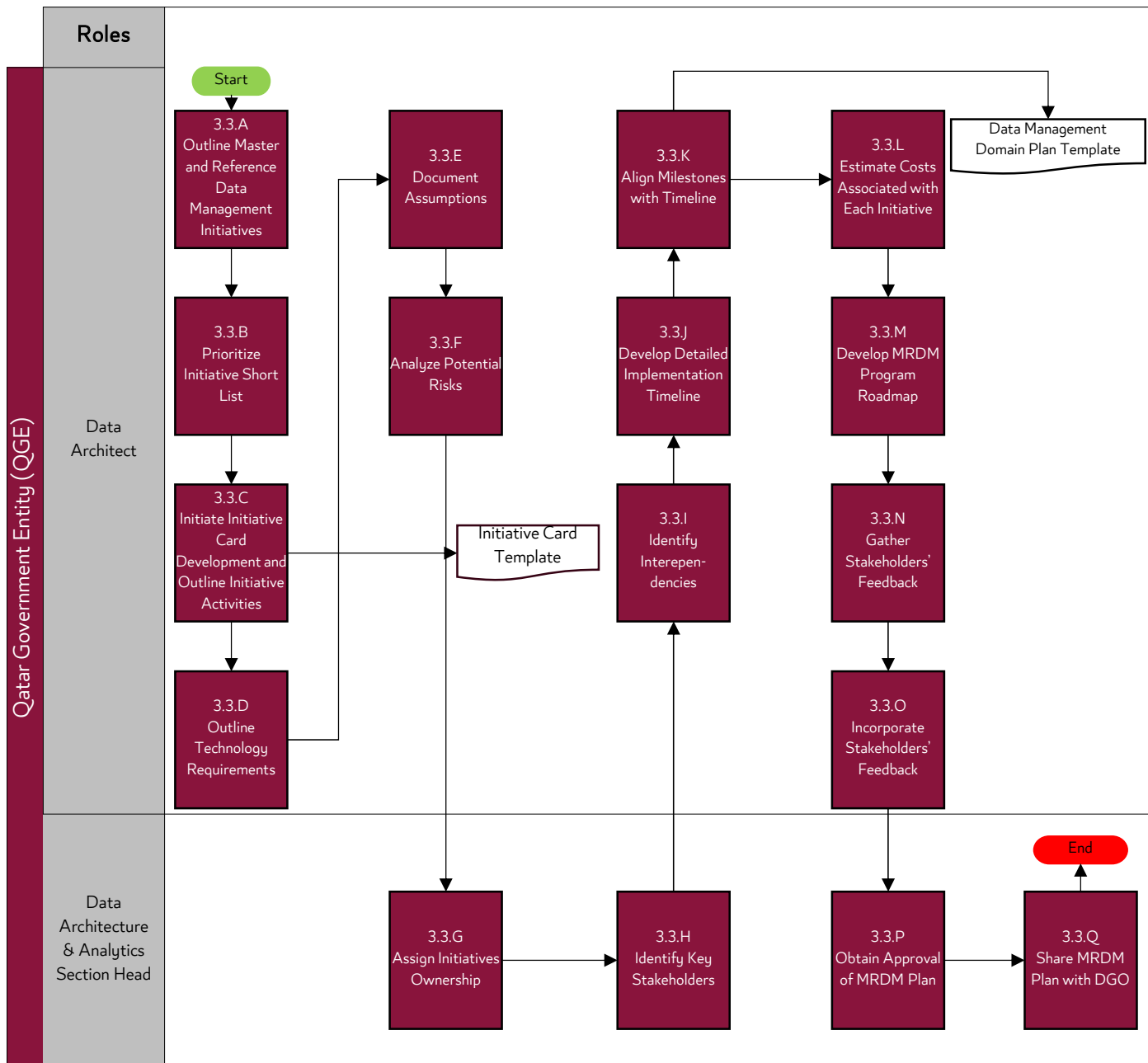


Figure 5: MRDM Plan Process Map

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

3.3.A

Outline Master and Reference Data Management Initiatives: Identify and document a long list of master and reference data management initiatives, encompassing both short-term and long-term initiatives. Initiatives may include efforts such as selecting the desired MDM Hub Architecture, setting MRD attribute trust, similarity, and survivorship rules, and selecting an MDM tool, as well as larger-scale projects like implementing the MDM Hub, and leveraging AI in the selected MDM Tool. Ensure these initiatives align with the Entity's broader data strategy to support effective resource allocation and address critical master and reference data needs.

3.3.B

Prioritize Initiative Short List: Leverage the master and reference data management initiatives long list to identify and prioritize a short list based on potential impact and feasibility. Focus on high-impact areas where master and reference data improvements will most significantly enhance decision-making and operational efficiency, ensuring that efforts support the Entity's strategic objectives. Additionally, identify quick wins—initiatives that can be implemented rapidly with minimal resources—to provide immediate benefits and demonstrate the value of effective master and reference data management.

3.3.C

Initiate Initiative Card Development and Outline Initiative Activities: Develop initiative cards for each prioritized initiative outlined in the short list, capturing essential activities and steps required to guide implementation. For each initiative, outline the key activities by breaking them down into specific steps and sub-activities that need to be completed in sequence (where applicable). Include descriptions of each step to provide clear guidance on execution, along with any necessary details on tools, techniques, or methodologies that may be required.

Asset Guidance

An **Initiative Card Template** provides a structured format to document key information related to prioritized initiatives, such as initiative overview, key activities, sponsor, owner, stakeholders, timeline, key milestones, budget, technology requirements, dependencies, assumptions, risks, and milestones. The initiative cards provide a structured overview of proposed projects, allowing stakeholders to quickly understand each initiative's purpose, expected impact, and resource requirements.

Illustrative

OVERVIEW xxx		SPONSOR xxx	STAKEHOLDERS xxx
		OWNER xxx	
KEY ACTIVITIES		TIMELINE	BUDGET (IN QAR)
Key Activity - Sub-activity X - Detailed Steps - Sub-activity Y - Detailed Steps Key Activity - Sub-activity X - Detailed Steps - Sub-activity Y - Detailed Steps		xxx	QAR xxx
		EXPECTED BENEFITS	
		• xxx	
		KEY MILESTONES	
		A. xxx B. xxx C. xxx	
		TECHNOLOGY REQUIREMENTS	
		• xxx	

OVERVIEW xxx		SPONSOR xxx	STAKEHOLDERS xxx											
		OWNER xxxxx												
KEY ACTIVITIES		TIMELINE												
#	Milestones & Key Activities	#	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
A.	Milestone	A.												
1.	Key Activity	1.												
B.	Milestone	B.												
2.	Key Activity	2.												
C.	Milestone	C.												
3.	Key Activity	3.												
DEPENDENCIES		ASSUMPTIONS & RISKS												
• xxx		• xxx												

3.3.D

Outline Technology Requirements: Identify and document the technology requirements necessary to support each prioritized MRDM initiative. This includes outlining the required tools, software, and systems needed to ensure successful implementation and adoption across the Entity.

3.3.E

Document Assumptions: Record assumptions including but not limited to resource availability, stakeholder engagement, and data complexity to ensure transparency and guide decision-making.

3.3.F

Analyze Potential Risks: Clearly identify and assess the risks associated with the master and reference data management initiatives. Examples include but are not limited to:

- **Limited Support for Standardization:** Low stakeholder buy-in for standardized data definitions and processes could hinder consistency and interoperability in master and reference data.
- **Data Privacy Compliance:** Failure to address privacy concerns early may result in non-compliance risks, particularly when handling sensitive data.
- **Insufficient Ownership and Stewardship:** Lack of defined data ownership and stewardship roles may weaken accountability and governance, affecting implementation success.
- **Dependence on External Data:** Heavy reliance on external data sources may pose risks if these sources are inconsistent, unavailable, or below entity standards.
- **Resource Shortages:** Specialized skills in data management and governance may be lacking, potentially delaying or compromising initiative success.

3.3.G

Assign Initiatives Ownership: The Data Management and Governance Section Head will assume ownership of all the master and reference data management initiatives. This designation ensures that all the initiatives have a dedicated leader who can oversee progress and resolve roadblocks that may arise.

3.3.H

Identify Key Stakeholders: Identify and engage key stakeholders critical to the respective MRDM initiatives, including data users, IT, data stewards, and management. Early involvement ensures alignment, clarifies responsibilities, and facilitates effective communication, supporting successful execution of the initiative.

3.3.I

Identify Interdependencies: Analyze how MRDM initiatives may impact or be influenced by planned initiatives or existing processes to improve resource allocation and coordination. This step involves:

- Reviewing how improvements in MRDM could affect or be affected by other business workflows.
- Identifying resource overlaps where the same teams, tools, or data sources are involved in multiple initiatives.
- Ensuring proper coordination between different departments to avoid conflicts or delays in resource allocation.

3.3.N

Gather Stakeholders' Feedback: Present the draft plan to all relevant stakeholders, inviting input on its objectives, scope, proposed activities, and overall timeline. This step ensures that any potential issues, concerns, or overlooked areas are identified early, improving the plan's alignment with organizational needs and building stakeholder commitment.

3.3.O

Incorporate Stakeholders' Feedback: Incorporate constructive feedback from stakeholders to refine the plan, addressing identified concerns and ensuring it aligns with stakeholder expectations. This step fine-tunes the plan details and strengthens its practical feasibility by leveraging diverse perspectives.

3.3.P

Obtain Approval of Master and Reference Data Management Plan: After integrating all stakeholder feedback, present the finalized plan to the Data Committee or relevant authority for formal approval. This step ensures that the plan aligns with organizational priorities and secures the necessary support for successful implementation. Once approved, the plan can be formally adopted and disseminated.

3.3.Q

Share Master and Reference Data Management Plan with DGO: Submit the finalized Master and Reference Data Management Plan to the National Data Governance Office (DGO) for review. The DGO will review the plan and provide recommendations, allowing the Entity to ensure their plan aligns with the broader national data program agenda and best practices.

RACI Matrix – MRDM Plan

Indicative

Sr No.	Activity	DGO	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Architect	Data Stewards	Data Strategy Expert
3.3.A	Outline Master and Reference Data Management Initiatives		A	C	C	R	C	C
3.3.B	Prioritize Initiative Short List		A	C	C	R	C	C
3.3.C	Initiate Initiative Card Development and Outline Initiative Activities			A	C	R		C
3.3.D	Outline Technology Requirements			A	C	R	C	C
3.3.E	Document Assumptions			A	C	R	C	C
3.3.F	Analyze Potential Risks			A	C	R		C
3.3.G	Assign Initiatives Ownership		A	R				
3.3.H	Identify Key Stakeholders		A	R	C	C	C	
3.3.I	Identify Interdependencies			A	C	R		C
3.3.J	Develop Detailed Implementation Timeline			A	C	R		C
3.3.K	Align Milestones with Timeline			A	C	R		C
3.3.L	Estimate Costs Associated with Each Initiative			A	C	R	C	C
3.3.M	Develop Master and Reference Data Management Plan Roadmap			A	C	R		C
3.3.N	Gather Stakeholders' Feedback			A	C	R		
3.3.O	Incorporate Stakeholders' Feedback			A		R		C
3.3.P	Obtain Approval of the Master and Reference Data Management Plan		A	R				
3.3.Q	Share Master and Reference Data Management Plan with DGO	C/I	A	R				

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Outline MRDM Initiatives



Develop and Review MRDM Plan

4. Master and Reference Data Management Assessment and Design



4. Master and Reference Data Management Assessment and Design

What is a Master and Reference Data Management Assessment and Design, and why does the Entity need it?

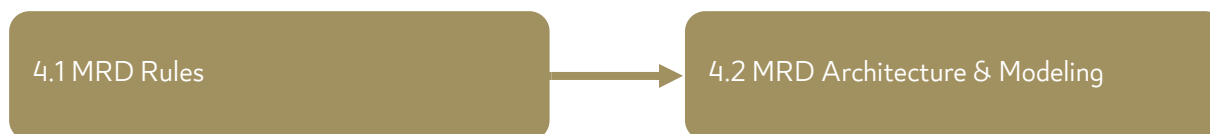
Master and Reference Data Management (MRDM) Assessment and Design is the process of establishing the rules and architectural framework required to manage master and reference data effectively. It involves defining attribute rules (e.g., trust, similarity, survivorship) for accurate matching, merging, and deduplication of data, as well as adopting common architectural archetypes (e.g., registry, consolidation, coexistence, centralized) to ensure scalability, interoperability, and data integrity.

The need for MRDM Assessment and Design arises from several critical factors:

1. **Data Quality and Consistency:** Ensures accurate, deduplicated, and consistent master and reference data by applying trust, similarity, and survivorship rules.
2. **Streamlined Data Governance:** Establishes clear rules for managing data attributes and governance processes to ensure data integrity across multiple systems.
3. **Efficient Data Management:** Supports the effective handling of data by adopting architectural archetypes that align with the Entity's needs and business goals.

Implementing MRDM Assessment and Design empowers Entities to streamline data governance, optimize data management processes, and support business operations with accurate and trustworthy data. It ensures data is consistently maintained and leveraged as an asset across the Entity.

The remaining sections of this chapter will detail the below activities that will guide the Entity on executing the below:



Guiding Principles for Master and Reference Data Management Assessment and Design

1. **Establish Attribute Rules:** Define and implement attribute trust, similarity, and survivorship rules to maintain accurate, consistent, and deduplicated data records.
2. **Standardize for Interoperability:** Ensure MDM hub architecture supports seamless integration and interoperability across departments and systems.
3. **Prioritize High-Value Data:** Focus on managing critical master and reference data assets that drive business operations and align with organizational priorities.
4. **Promote Continuous Improvement:** Regularly review attribute rules and architectural models to adapt to evolving business needs and technological advancements.

4.1 MRD Rules

The success of a data management program hinges on the ability to standardize, deduplicate, and unify critical data across the Entity. Match and merge rules serve as the cornerstone for this unification process, enabling Entities to identify duplicates, resolve discrepancies, and consolidate data into a single, authoritative "golden record."

Attribute trust, similarity, and survivorship rules form the foundation of the match and merge process. These rules dictate how records are compared, evaluated, and resolved during the data consolidation lifecycle. While match and merge provide the overarching framework for identifying and resolving duplicate records, the attribute-based rules define the detailed mechanics for:

1. **Assessing trustworthiness** of data attributes from different sources.
2. **Quantifying similarity** between attributes to identify potential matches.
3. **Resolving conflicts** through survivorship rules to select the most accurate, reliable, or appropriate data for the golden record.

By establishing clear and consistent attribute trust, similarity, and survivorship rules, entities can ensure the accuracy, integrity, and reliability of their master and reference data.

This activity will outline the principles behind defining these rules and their role in enabling effective match and merge operations.

4.1.1 Define Attribute Trust Rules

This activity outlines the steps to define and implement Attribute Trust Rules, essential for ensuring the reliability and accuracy of data attributes across master and reference data domains. This activity enables the Entity to create a robust system to assess and maintain trust in its CDEs.

Process – Define Attribute Trust Rules

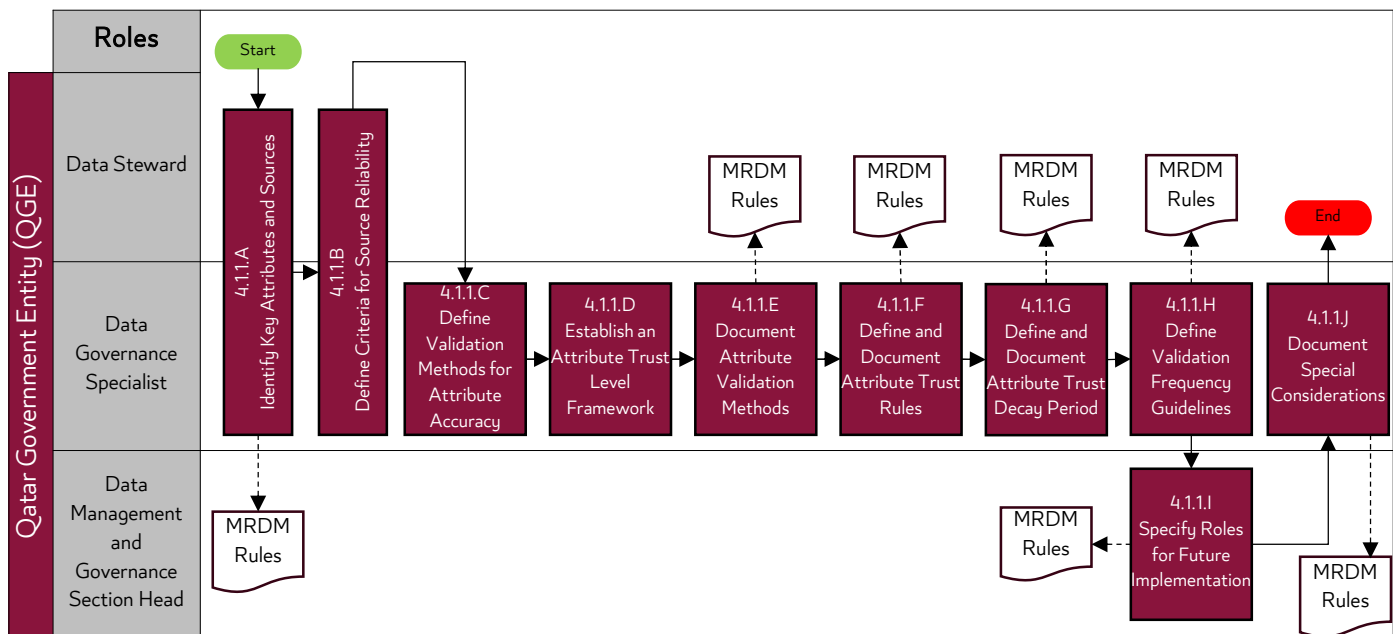


Figure 6: Process Map for Defining Attribute Trust Rules

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

4.1.1.A

Identify Key Attributes and Sources: Reference the Data Catalog and review the prioritized CDEs and their respective mapped sources to ensure the CDEs cover all relevant MRD domains (e.g., customer, product, supplier). Highlight attributes with multiple sources to identify potential overlaps or conflicts that may need resolution during the trust definition process.

Populate the **Attribute** and **Data Source** columns within the **MRDM Rules Template** for the prioritized CDEs and respective data sources. This step sets the foundation for applying trust rules by confirming that the prioritized CDEs and their sources are accurate and comprehensive.



Key Message / Information

Before proceeding, the Entity must ensure that its Data Catalog is fully populated, as this step leverages the prioritized CDEs (attributes) outlined in the Data Catalog that has been developed in alignment with the Entity's data strategy and Data Catalog and Metadata Management domain. These CDEs represent the attributes deemed most critical to the Entity's operations, decision-making, or regulatory compliance. Since the Data Catalog includes mappings of attributes to their respective source systems, the focus here is to confirm and validate the completeness of these mappings and ensure alignment with broader master and reference data management objectives.



Asset Guidance

The **MRDM Rules Template** contains the **Attributes Trust Rules Sheet** which the Entity can use to document prioritized CDEs, respective data source, source update frequency and historical accuracy, validation methods, rules, and frequency, and decay period.

Illustrative

Master and Reference Data Management Attribute Trust Rules									
The following table provides a structured template to capture MRD trust rules and manage the trustworthiness of MRD attributes. Use this table to document the CDEs and define the respective data sources, source update frequency, source historical accuracy, technical validation methods, technical validation rules, and required frequency to perform each validation method to ensure the accuracy and reliability of key data attributes within the Entity. Regular validation helps maintain high data quality and consistency.									
Field	Description								
Attribute	The specific data attribute (e.g., Customer Name, Phone Number) for which trust rules are applied.								
Data Source	The source system where the data originates as indicated in the Data Catalog (e.g., CRM, ERP, external sources).								
Source Update Frequency	The frequency in which the data source is being updated.								
Source Historical Accuracy	The source's track record of providing accurate data over time.								
Technical Validation Method	The method used to check the reliability of the data (e.g., cross-referencing, domain-specific range checks, business rule validation etc.)								
Technical Validation Rule Description	The description of the attribute validation check (e.g., check email address for syntactic correctness (e.g., presence of @domain.com))								
Validation Frequency	How frequently the technical validation methods should be executed for the respective attribute.								
Trust Decay Period	Length of time that the trust level takes to decay from the maximum trust to the minimum trust								
Responsible Party	The role or team responsible for managing the trust rules and validating the data.								
Comments	Additional information or special considerations for this attribute.								
Attribute	Data Source	Source Update Frequency	Source Historical Accuracy	Technical Validation Method	Technical Validation Rule Description	Validation Frequency	Trust Decay Period	Responsible Party	Comments



Key Message / Information

In the instance that there are no overlapping data sources for a specific data attribute, there is no need to define attribute trust rules.

4.1.1.B

Define Criteria for Attribute Source Reliability: Establish criteria to evaluate the reliability of the key attributes respective data sources. This involves assessing key factors including but not limited to:

- **Technical Validation:** The methods of technical validation that will be applied during the attribute trust assessment to ensure attribute data quality dimension results are satisfactory. Strong validation practices enhance data reliability.
- **Source Update Frequency:** How often the data is updated in the source system (e.g., real-time, daily, monthly). Frequent updates indicate that the source is actively maintained and more likely to provide current information.
- **Source Historical Accuracy:** The source's track record of providing accurate data over time. Sources with low errors and minimal corrections are more trustworthy.

By evaluating these factors, the Entity can create a foundation for assessing the trustworthiness of the sources supplying the data attributes.

4.1.1.C

Define Technical Validation Methods for Attribute Accuracy: Building upon the **Technical Validation** criterion defined in the previous step, identify and define the specific technical validation methods that will be used to verify attribute data quality dimensions (e.g., accuracy, consistency, and completeness etc.) during the implementation phase. Outline a clear explanation **of how each technical validation method influences attribute trust levels**. Common technical validation methods include but are not limited to:

Cross-Referencing with Authoritative Systems:

- Compare attributes against trusted authoritative systems (e.g., gov. registries).
- **Example:** Validate National ID Number against a government database or Business Registration Number with a trade registry. Attributes verified through such cross-referencing earn higher trust scores due to their reliance on authoritative sources.

Data Profiling:

- Analyze the internal consistency, completeness, and format of attribute values to detect anomalies.
- **Example:** Check Email Address for syntactic correctness (e.g., presence of @domain.com) and Customer Phone Number for proper country code formatting. While profiling doesn't confirm correctness, it flags inconsistencies for further review.

Domain-Specific Range Checks:

- Validate numeric or categorical attributes to ensure they fall within predefined ranges or match accepted values.
- **Example:** Validate that Age is within the allowed range for a service (e.g., 18-99 for adult services) or ensure that Product Categories belong to an approved taxonomy.

Pattern and Format Validation:

- Ensure that attributes adhere to predefined patterns or formats using validation rules or regular expressions (regex).
- **Example:** Validate Customer Phone Number to follow the E164 format or check Order IDs for adherence to a consistent alphanumeric structure (e.g., "ORD-XXXXX").

4.1.1.C
Cont.**Business Rule Validation:**

- Validate attributes based on predefined business logic or constraints.
- **Example:** Ensure that Order Amount is always greater than Discount Amount or that Invoice Date is before the Due Date. Adherence to these rules demonstrates logical consistency and boosts trust in the data.

4.1.1.D

Establish an Attribute Trust Level Framework: Once the criteria for attribute source reliability (source criteria) and the respective validation methods are defined, design a framework to assess attribute trust levels. The framework should outline what makes an attribute trustworthy by assigning validation methods within each source criterion to trust categories. Sample categories include:

- **Numerical:** Assign technical validation methods (e.g., regex validation, authoritative external validation) to specific levels of numerical trustworthiness (e.g., 1-10).
- **Categorical:** Group validation methods into trust categories (e.g., Low, Med, High).
- **Hybrid:** Assign validation methods to specific levels of numerical trustworthiness and group numerical trustworthiness by overarching categorical trustworthiness.

For example, data with **frequent updates** and **automated external validation (e.g., API)** may receive **higher trust levels** than data with **infrequent updates** and **basic regex checks** (e.g., validating an email with "@domain.com"). An illustrative trust matrix using the hybrid approach is shown below:

Source Criteria	Validation Methods (For Each Source Criteria)		
	Low Trust (1-3)	Medium Trust (4-7)	High Trust (8-10)
Technical Validation	Regex validation for basic syntax	Internal peer-to-peer cross-checks	Authoritative external validation (e.g., API)
Source Update Frequency	Infrequent (quarterly or longer)	Regular updates (weekly or monthly)	Frequent updates (real-time or daily)
Source Historical Accuracy	High error rates or inconsistent corrections	Moderate errors rate, mostly consistent	Rare errors, consistently accurate over time

4.1.1.E

Document Attribute Validation Methods: Populate the **Source Update Frequency**, **Source Historical Accuracy**, and **Technical Validation Method** columns within the **MRDM Rules Template** with the appropriate validation methods outlined in the **Attribute Trust Level Framework (4.1.1.D)**. These methods are selected based on their ability to verify attribute data quality dimensions (e.g., accuracy, consistency, and completeness etc.). Each validation method should reflect the rigor and reliability needed to assess trustworthiness for the attribute.

**Key Message / Information**

The choice of validation method depends on the attribute's importance, criticality, and the availability of authoritative or internal systems for comparison. For example, cross-referencing with authoritative systems provides a high level of trust for attributes like National ID or Tax Number, while methods like format validation or data profiling may be sufficient for attributes like Phone Number or Email Address.

4.1.1.F

Define and Document Attribute Trust Rules: Define the attribute trust rules (**Technical Validation Rule Description**) for the attributes and technical validation methods previously identified. These rules specify the checks and logic applied to verify data quality dimensions

Populate the **Technical Rule Description** column in the **MRDM Rules Template** with details of the validation checks for each attribute. This includes the logic, thresholds, or conditions under which the technical validation methods will be executed to assess trustworthiness.



Key Message / Information

Defining attribute trust rules ensures a consistent and measurable approach to evaluating attribute trustworthiness. Example attribute trust rules for Phone Number could include performing format validation to ensure compliance with the E.164 standard or ensuring that an Order Amount is always greater than the Discount Amount.

4.1.1.G

Define and Document Attribute Trust Decay Period: For each attribute, considering the respective data source, determine how long the trust level of the attribute remains valid before it starts to diminish. The trust decay period reflects the frequency of updates, the nature of the attribute, and its importance to operational processes.

Populate the **Trust Decay Period** column in the **MRDM Rules Template** with the defined time period (e.g., 3 months, 6 months, 1 year) for each attribute. This period ensures that trust levels remain dynamic and responsive to changes in data quality or updates.



Key Message / Information

Trust decay periods help ensure that outdated or unverified attribute values are identified and reassessed regularly. For example, Email Address or Contact Number may have shorter trust decay periods (e.g., 6 months) due to frequent updates, while attributes like National ID or Tax Number may have longer decay periods (e.g., 2 years) because they are more stable. By defining trust decay periods, the Entity can maintain ongoing confidence in the trust levels of critical attributes.

4.1.1.H

Define Validation Frequency Guidelines: Propose validation frequency guidelines based on the criticality of each attribute. For instance, customer contact information (e.g., Email Address) may require monthly validation due to its frequent usage in communication, while product codes or inventory data might only need quarterly or annual checks. This ensures that validation efforts are proportionate to the importance and usage of the attributes.

Populate the **Validation Frequency** column within the **MRDM Rules Template** with the recommended schedule for verifying attribute accuracy (e.g., monthly, quarterly, annually).

4.1.1.I

Specify Roles for Future Implementation: Identify and document the roles responsible for maintaining trust levels during implementation, periodic reviews, and the reporting of issues or anomalies. While the actual assignment of trust levels will occur later in the implementation phase, it is important to define accountability in advance. Typical roles may include Data Stewards, Data Governance experts, or domain experts who will oversee the application and ongoing maintenance of trust levels.

4.1.1.J

Document Special Considerations: Finally, document any special considerations or exceptions for specific attributes. This may include attributes sourced from third-party systems, attributes requiring manual overrides, or attributes subject to additional regulatory or business requirements. Capturing these nuances ensures that the trust rules framework accommodates all unique scenarios.

RACI Matrix – Define Attribute Trust Rules

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Steward	Metadata Specialist	Data Owner
4.1.1.A	Identify Key Attributes and Sources	I	A	R	R	C	C
4.1.1.B	Define Criteria for Source Reliability		A	R	R		C
4.1.1.C	Define Validation Methods for Attribute Accuracy		A	R	C		C
4.1.1.D	Establish an Attribute Trust Level Framework		A	R	C		
4.1.1.E	Document Attribute Validation Methods		A	R	C		
4.1.1.F	Define and Document Attribute Trust Rules		A	R	C		
4.1.1.G	Define and Document Attribute Trust Decay Period		A	R	C		
4.1.1.H	Define Validation Frequency Guidelines		A	R	C		
4.1.1.I	Specify Roles for Future Implementation	A	R	I			
4.1.1.J	Document Special Considerations		A	R	C	C	

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Define and Document Attribute Trust Rules for CDEs

4.1.2 Define Attribute Similarity Rules

This activity provides a structured approach for defining and implementing Attribute Similarity Rules, essential for accurate deduplication and matching processes across datasets. The process ensures that potential duplicates or related records are consistently and effectively identified, enhancing data quality and reliability.

Process – Define Attribute Similarity Rules

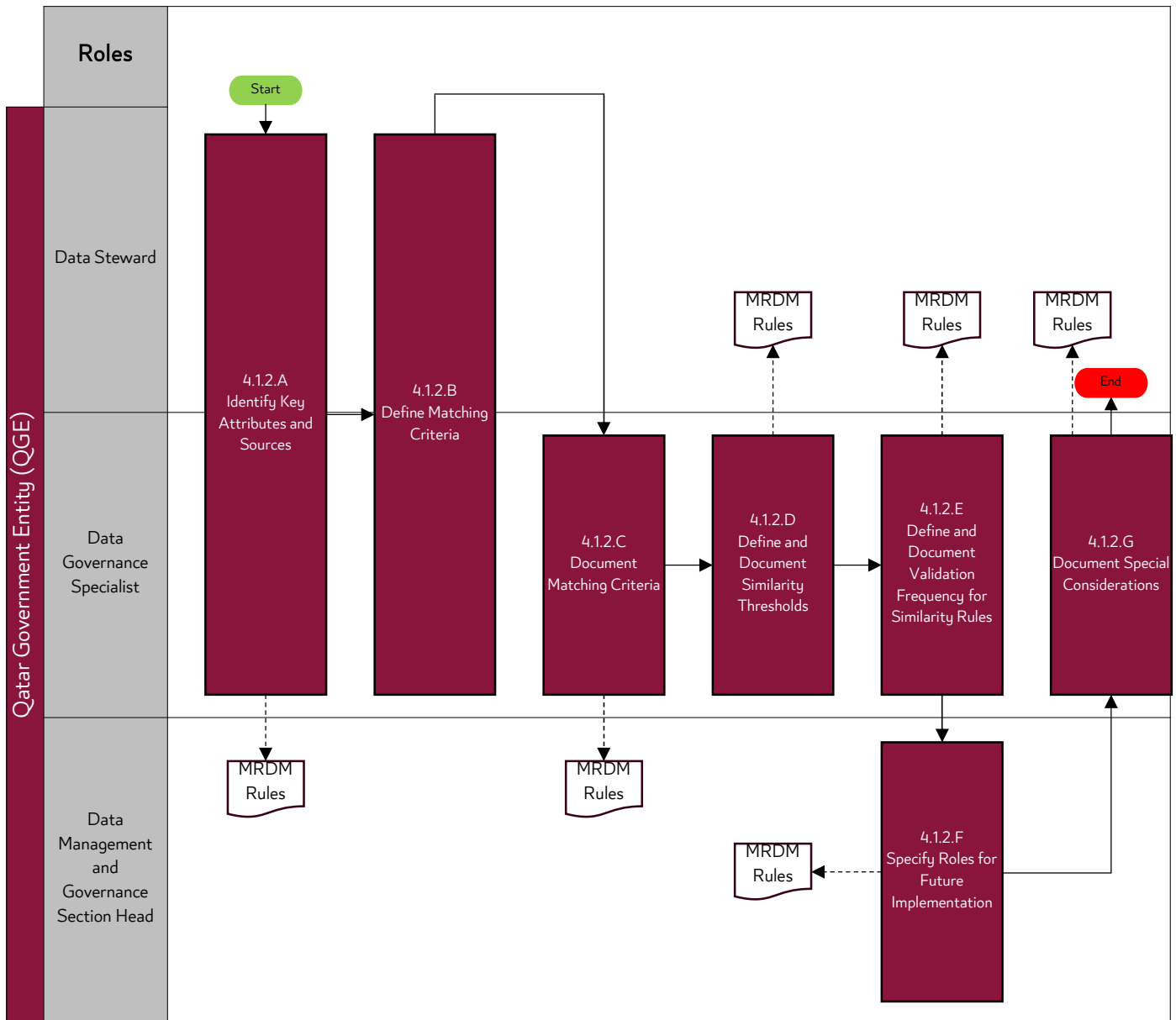


Figure 7: Process Map for Defining Attribute Similarity Rules

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

4.1.2.A

Identify Key Attributes and Sources: Reference the Data Catalog and review the prioritized CDEs to identify attributes that require similarity rules for the deduplication or matching process. These are often attributes prone to variation across datasets, such as Customer Name, Email Address, or Phone Number.



Tip

- Focus on attributes where inconsistencies (e.g., typos, different formats) may lead to mismatches without similarity rules.
- Highlight attributes shared across multiple datasets to ensure consistency during comparison.



Key Message / Information

Attributes requiring similarity rules often involve free-text fields or fields with format variability. For example, names may include variations (Mohammed H. Mohsen vs. M. Mohsen), and addresses may differ due to abbreviations (St. vs. Street). This step confirms the scope of attributes that require similarity logic.

Populate the **Attribute** and **Data Source** column within the **MRDM Rules Template** with the prioritized CDEs and their respective data sources. This step ensures that all attributes requiring similarity rules are clearly identified and mapped to their original source systems. By doing so, the foundation is established for accurately comparing attributes across datasets, enabling effective deduplication or matching processes.

Asset Guidance

The **MRDM Rules Template** contains the **Attributes Similarity Rules Sheet** which the Entity can leverage to document attribute matching criteria, techniques, thresholds, and actions on the respective match for prioritized CDEs.

Illustrative



Master and Reference Data Management Attribute Similarity Rules

The following table provides a structured template to define the criteria and methods for identifying similar records across master and reference data. This helps in preventing duplicates and ensuring that matching records are appropriately merged or flagged for review. Use this table to specify matching techniques, thresholds, and actions to be taken on potential matches.

Field	Description
Attribute	The data attribute being compared for similarity (e.g., Email Address, Phone Number).
Data Source	The source system where the data originates as indicated in the Data Catalog (e.g., CRM, ERP, external sources).
Matching Criteria Type	The type of matching criteria used for comparing records (e.g., fuzzy match, exact match).
Matching Criteria	The defined logic used to determine data matches, specifying the exact criteria applied (e.g., "QID number must match exactly" or "DOB field must be within ±1 day")
Threshold (%)	The similarity threshold (in percentage) used to determine when two records are considered a match.
Matching Technique	The algorithm or method used for similarity matching (e.g., Exact Match, Levenshtein distance, Soundex).
Validation Frequency	How frequently the similarity (matching) rules should be executed for the respective attribute.
Responsible Party	The team or individual responsible for reviewing and approving matches based on the rules.
Comments	Additional notes or guidelines for special cases or exceptions in similarity matching.

Attribute	Data Source	Matching Criteria Type	Matching Criteria	Threshold (%)	Validation Frequency	Responsible Party	Comments

4.1.2.B

Define Matching Criteria: Select the appropriate matching criteria for each attribute to ensure accurate and effective comparison of records. Matching criteria define how two or more records are compared, guiding the match engine in determining whether they are the same or similar enough to be considered a match. Different attributes may require different criteria depending on their type, format, and variability across datasets. Common matching criteria include but are not limited to:

Exact Match:

- Compares attribute values directly to determine if they are identical.
- **Use Case:** Ideal for highly structured attributes where variations are minimal, such as National ID, Social Security Number, or Employee ID.
- **Example:** Comparing 12345678 with 12345678 results in a match; 12345678 and 12345679 do not match.

Fuzzy Match:

- Uses algorithms to measure the similarity between two strings, allowing for minor variations or typographical errors.
- **Use Case:** Suitable for names (Mohammed H. Mohsen vs. Mohammad Mohsen), addresses, or other free-text fields.
- **Example:** Matching Ahmad and Ahmed with a high similarity score. Algorithms include Levenshtein Distance, Jaro-Winkler, or Cosine Similarity.

Range Match:

- Matches numeric or date attributes within a specified range.
- **Use Case:** Appropriate for Date of Birth (e.g., allowing for errors in data entry) or Transaction Amount (e.g., matching amounts within a tolerance).
- **Example:** Match 2023-01-01 with 2023-01-02 if the allowable range is ± 1 day.

Pattern Match:

- Matches attributes using defined patterns or regular expressions.
- **Use Case:** Effective for attributes with specific formats like Phone Numbers, Email Addresses, or Order IDs.
- **Example:** Match ORD-12345 with ORD-12345 by identifying a consistent pattern of ORD-XXXXX.

Tokenized Match:

- Splits attribute values into smaller components (tokens) and matches them individually.
- **Use Case:** Suitable for long fields like Address or Full Name, where reordering or partial matches are common.
- **Example:** Matching Mohammed H. Mohsen with Mohsen, Mohammed.

Contextual Match:

- Matches records based on contextual rules or additional supporting data.
- **Use Case:** Effective for Addresses or Geographic Locations where multiple data points (e.g., city, state, and postal code) influence the match decision.
- **Example:** Match 123 Main St., Springfield with 123 Main Street, Springfield by confirming the city and postal code align.

4.1.2.C

Document Matching Criteria: Populate the **Matching Criteria Type** and **Matching Criteria** columns within the **MRDM Rules Template** by specifying how each attribute will be compared. Each matching criterion should be tailored to the attribute's type, variability, and purpose. For instance, an exact match may be appropriate for structured attributes like Employee ID, while fuzzy matching is better suited for variable attributes like Employee Name.

By focusing on attribute-specific matching criteria types and matching criteria logic that account for data patterns and common discrepancies, entities can enhance the accuracy and reliability of their similarity rules. This ensures that potential duplicates or related records are identified consistently and effectively across datasets.



Key Message / Information

When implementing similarity rules for data attributes, it is essential to first ensure that the attributes are in a comparable format. For example, when comparing phone numbers in a database, it is important to standardize their data type to avoid mismatches, such as one phone number being stored as an integer and another stored as a string.

4.1.2.D

Define and Document Similarity Thresholds: Define and document similarity thresholds for each attribute in the **Threshold (%)** column of the **MRDM Rules Template** to specify when two values are considered a match. Indicative thresholds are seen below:

- **High Similarity Threshold (e.g., >90%):** Indicates a strong match; often used for exact or near-exact matches.
- **Medium Similarity Threshold (e.g., 70%-90%):** Indicates a potential match requiring manual review.
- **Low Similarity Threshold (<70%):** Indicates values are unlikely to be matches.



Key Message / Information

Thresholds should be tailored based on the attribute type and use case. For instance, a high threshold is appropriate for numeric identifiers, while a medium threshold may suffice for names where minor variations are common.

4.1.2.E

Define and Document Validation Frequency for Similarity Rules: Propose validation frequency guidelines based on the criticality and likelihood of changes for each attribute. Populate the **Validation Frequency** column within the **MRDM Rules Template** with the recommended schedule for reviewing and updating similarity rules (e.g., monthly, quarterly, annually).

4.1.2.F

Specify Roles for Future Implementation: Identify and document the roles responsible for maintaining the similarity rules during implementation, periodic reviews, and reporting of issues. While the actual implementation will occur later in the implementation phase, defining accountability in advance reduces the risk of mismatches. Typical roles include Data Stewards, Data Governance experts, or domain experts who oversee the application and maintenance of similarity rules.

4.1.2.G

Document Special Considerations: Document any special considerations or exceptions for specific attributes. This may include attributes with inconsistent formats across systems (e.g., Date of Birth stored in different date formats), attributes requiring custom similarity thresholds, or attributes impacted by unique regulatory requirements. Capturing these nuances ensures the approach to defining similarity rules is tailored to handle unique attribute requirements and edge cases.

RACI Matrix – Define Attribute Similarity Rules

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Steward	Metadata Specialist	Data Owner
4.1.2.A	Identify Key Attributes and Sources	I	A	R	R	C	C
4.1.2.B	Define Matching Criteria	I	A	R	R		C
4.1.2.C	Document Matching Criteria	I	A	R	C		
4.1.2.D	Define and Document Similarity Thresholds	I	A	R	C		
4.1.2.E	Define and Document Validation Frequency for Similarity Rules	I	A	R	C		
4.1.2.F	Specify Roles for Future Implementation	A	R	I			
4.1.2.G	Document Special Considerations		A	R	C		

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Define and Document Attribute Similarity Rules

4.1.3 Define Attribute Survivorship Rules

This activity outlines a structured approach for defining Attribute Survivorship Rules, essential for resolving conflicts between attributes and determining which attribute values are retained when consolidating duplicate records into a single, authoritative “golden record”. Survivorship rules ensure data accuracy, reliability, and compliance by specifying retention criteria and addressing edge cases.

Process – Define Attribute Survivorship Rules

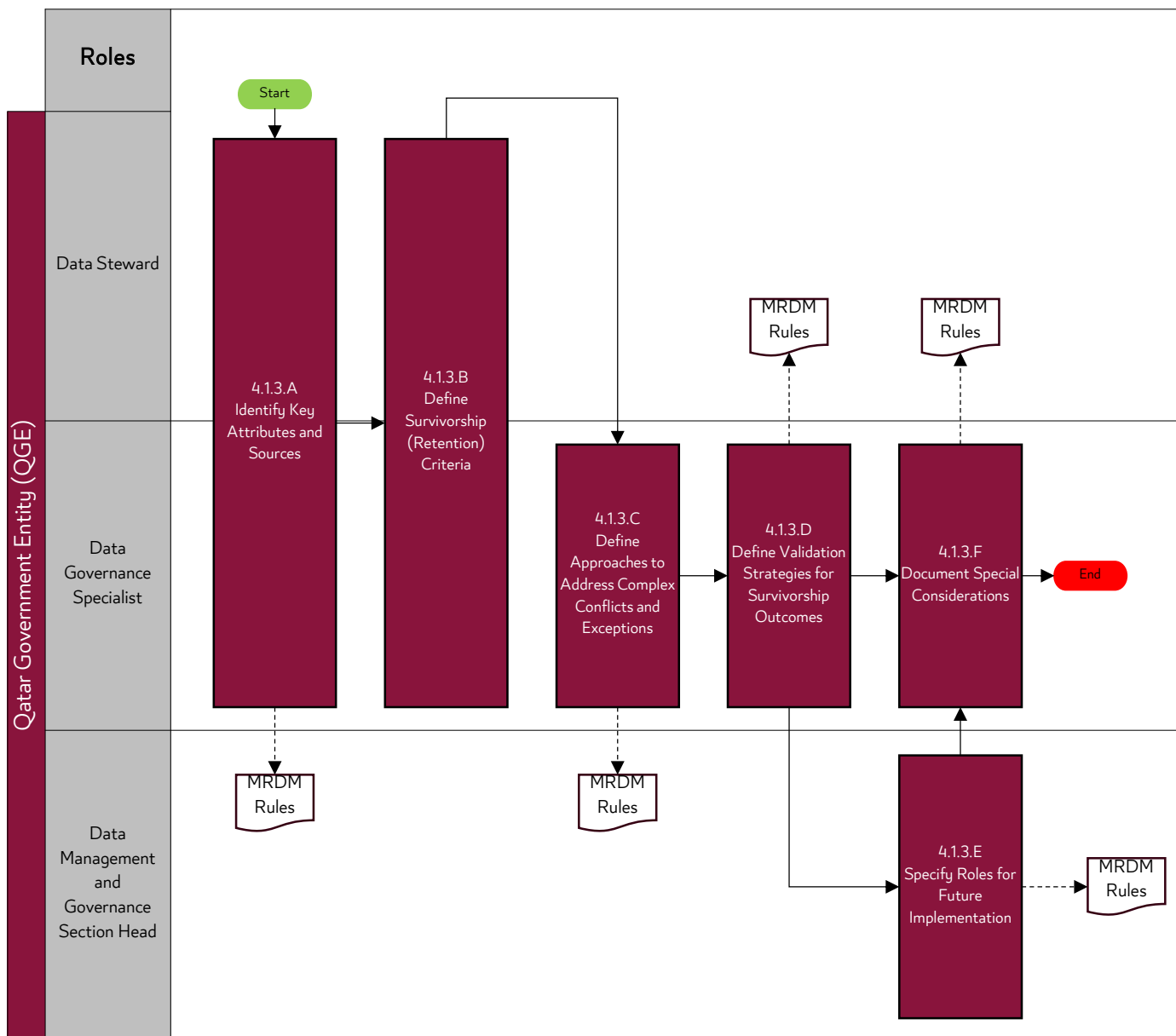


Figure 8: Attribute Survivorship Rules Process Map

Legend: ■ Process Step ◆ Decision ● Process Start ● Process End □ Document ■ Off-Page Reference

4.1.3.A

Identify Key Attributes and Sources: Reference the Data Catalog and review the prioritized CDEs to identify attributes that are commonly duplicated across records and require survivorship logic. These include attributes sourced from multiple systems or frequently updated fields where consistency and accuracy are critical. Examples include but are not limited to:

- **Dynamic Attributes:** Phone Number, Email Address, or Last Modified Date.
- **Critical Identifiers:** Government ID, Tax Number, or Employee ID.
- **Context-Specific Attributes:** Residential Address or Contact Preferences.

Populate the **Attribute Name** column in the **MRDM Rules Template** with these prioritized attributes.



Key Message / Information

Attributes requiring survivorship rules are those prone to variation across systems or duplicates, where it is crucial to consolidate them accurately. For instance, a citizen's Residential Address might differ between databases, requiring rules to resolve conflicts.

Asset Guidance

The **MRDM Rules Template** contains the **Attributes Survivorship Rules Sheet** which the Entity can leverage to document attribute survivorship rule type, survivorship (retention) criteria, conflict resolution strategies, the preferred data source to use in case of conflicting values (primary source), the backup source if the primary source is unavailable or lacks data (secondary source), as well as the validation frequency and responsible party.

Illustrative

 Master and Reference Data Management Attribute Survivorship Rules	
The following table provides a structured template to define how conflicts between duplicate or multiple data records are resolved. Use this table to outline the rules for determining which data should 'survive' and be retained as the golden record, based on trusted sources, most recent data, or other tie-breaker rules. This ensures consistency and accuracy in your master data management.	

Field	Description
Attribute Name	The attribute where multiple conflicting values exist (e.g., Email Address, Date of Birth).
Survivorship Rule Type	The rule for deciding which attribute data should survive (e.g., most recent, highest trust score, custom business rule).
Survivorship (Retention) Criteria	A description of the specific survivorship (retention) criteria for the respective Survivorship Rule Type. E.g., For citizen academic qualifications, use logic to retain the the highest academic qualification verified by an official body or institution, over lower academic qualifications with unverified records.
Conflict Resolution	Strategies to address complex conflicts and exceptions where standard retention criteria cannot resolve discrepancies. E.g., source hierarchy prioritization, fallback rules, field-level validation etc.
Primary Source	The preferred data source to use in case of conflicting values (e.g., CRM, HR System).
Secondary Source	The backup data source if the primary source is unavailable or lacks data.
Validation Frequency	How frequently the survivorship rules should be executed for the respective attribute.
Responsible Party	The person or team responsible for implementing and managing survivorship decisions.
Comments	Additional information or exceptions to the survivorship rules.

Attribute Name	Survivorship Rule Type	Survivorship (Retention) Criteria	Conflict Resolution	Primary Source (Optional)	Secondary Source (Optional)	Validation Frequency	Responsible Party	Comments

4.1.3.B

Define Survivorship (Retention) Criteria: Establish clear criteria to determine which attribute value is retained when duplicate records are merged. Retention criteria ensure that the most accurate, reliable, and relevant data is selected for the golden record. The choice of criteria depends on the attribute's purpose, importance, and variability. Common criteria include but are not limited to:

Most Recent Value:

- Retain the attribute value with the latest timestamp.
- **Example:** For Residential Address, prioritize the address updated most recently, as citizens often change their addresses and only the latest address is valid for correspondence or services.

Highest Trust Source:

- Retain the value from the most authoritative or reliable source system.
- **Example:** For National ID Number, select the value from the official national registry database over other systems (e.g., CRM or ERP).

Most Frequent Value:

- Select the value that appears most often across duplicate records.
- **Example:** For Language Preference, use the most common value when multiple duplicates list different preferences, ensuring alignment with citizen interactions.

Longest or Most Complete Value:

- Retain the value with the most detail or completeness, particularly for free-text fields.
- **Example:** For Full Name, prioritize Abdullah Mohammed Al-Thani over A. Al-Thani, as the longer value is more complete and useful for official documents.

Domain-Specific or Custom Business Rule:

- Apply predefined rules or logic tailored to the attribute's purpose.
- **Example:** For Tax Filing Status, use logic to retain the "Approved" status from the Tax Department over pending or rejected statuses in secondary systems.

Latest Verified Value:

- Retain the value most recently verified by a trusted process or system.
- **Example:** For Phone Number, keep the number that was most recently verified through an SMS confirmation process.

Hierarchical Rules for Aggregated Data:

- Use a hierarchical approach for attributes aggregated from multiple records.
- **Example:** For Case Status in citizen complaints, prioritize values like "Resolved" or "Closed" over "In Progress" when merging multiple records for the same case.

Populate the **Survivorship Rule Type** and **Retention Criteria** column in the **Survivorship Rules Template** with the appropriate logic for each attribute.

**Key Message / Information**

Retention criteria must align with the attribute's role in government operations. For example, retaining the most recent value may suit dynamic attributes like Residential Address, while prioritizing authoritative sources is essential for Government ID.

4.1.3.C

Define Approaches to Address Complex Conflicts and Exceptions: Define strategies to address complex conflicts and exceptions where standard retention criteria cannot resolve discrepancies. Survivorship rules must handle cases where multiple attribute values meet the criteria or require additional logic to resolve conflicts. Common approaches include but are not limited to:

Source Hierarchy Prioritization:

- Establish a hierarchy of trusted source systems to determine which value is retained during consolidation.
- **Example:** For Citizen Name, prioritize the value from the Civil Registry (as the authoritative source) over CRM systems. If conflicts persist across sources, escalate for manual review to ensure the golden record reflects the most accurate value.

Fallback Rules:

- Define fallback values or default rules for attributes when no value meets the survivorship criteria.
- **Example:** For Tax ID, if no valid value exists across duplicate records, assign a placeholder such as "Not Verified" and flag the record for follow-up verification. This ensures no data is lost while highlighting records requiring additional action.

Manual Review Triggers:

- Identify specific scenarios where automated survivorship rules cannot resolve conflicts and require human intervention.
- **Example:** For Land Ownership Records, trigger manual review if duplicate records list conflicting owners. This ensures accuracy for legal and regulatory compliance, where errors could lead to disputes or enforcement challenges.

Weighted Scoring Models:

- Assign weights to data quality factors (e.g., source reliability, last updated date, and data completeness) to calculate a score for each value. Retain the value with the highest score.
- **Example:** For Employment Status, use a weighted scoring model that assigns higher weights to data verified by the Ministry of Labor, followed by the recency of the status update (e.g., within the last six months). Retain the status with the highest score for the golden record.

Field-Level Validation:

- Apply validation checks to prioritize the most reliable value for survivorship.
- **Example:** For Email Address, retain the value that passes domain validation (e.g., ensuring it is a real domain) and has been recently verified via citizen confirmation (e.g., through an OTP process). This ensures the golden record only includes active, valid email addresses.

Populate the **Conflict Resolution** column (and the **Primary Source** and **Secondary Source** columns if applicable) in the **MRDM Rules Template** with specific rules or fallback strategies for each attribute.



Key Message / Information

Complex conflicts require well-defined survivorship rules to ensure data accuracy and consistency across systems. For example, hierarchical prioritization may be most effective for Citizen Name, while weighted scoring is better suited for resolving Residential Address conflicts. These approaches address exceptions and edge cases systematically to create reliable golden records.

4.1.3.D

Define Validation Strategies for Survivorship Outcomes: Define an approach to validate the outcomes of survivorship rules once implemented and operationalized in the implementation phase. Validation ensures that the survivorship rules are effective in producing accurate and reliable golden records

Sample-Based Validation:

- Define a process for periodically reviewing a sample of merged records to confirm that survivorship rules are selecting the most appropriate attribute values.
- **Example:** Review merged citizen records to ensure the Residential Address reflects the latest verified address from the Ministry of Municipality.

Rule Effectiveness Metrics:

- Identify key metrics to evaluate the performance of survivorship rules. Metrics may include the percentage of records flagged for manual review, the accuracy of retained values based on authoritative sources, or error rates in the golden record.
- Example: Track the accuracy of Employment Status in merged records by comparing retained values with Ministry of Labor data.

Feedback Mechanism:

- Plan for user feedback or input from data stewards to identify discrepancies or areas where survivorship rules may need refinement.
- Example: Data stewards can flag inconsistencies in Land Ownership Records where conflicting ownership details require rule adjustments.

Validation Frequency Guidelines:

- Propose how often survivorship rules should be validated and updated (e.g., quarterly, annually) based on attribute criticality and system updates.
- Example: Validate Tax ID survivorship rules annually to align with updates from the General Tax Authority.

Populate the **Validation Frequency** column in the **MRDM Rules Template** with the planned approaches for validating survivorship rule outcomes.

**Key Message / Information**

It is essential to maintain the history of matches and merges so that they can be undone when discovered to be incorrect. Match and Merge rate metrics enable the Entity to monitor the impact and effectiveness of matching inference rules and merging rules.

4.1.3.E

Specify Roles for Future Implementation: Identify and document the roles responsible for maintaining survivorship rules during implementation, periodic reviews, and reporting of issues. While the actual implementation will occur later, defining accountability in advance reduces the risk of mismatches. Typical roles include Data Stewards, Data Governance experts, or domain experts who oversee the application and maintenance of survivorship rules.

4.1.3.F

Document *Special Considerations*: Capture any special considerations or exceptions for specific attributes to ensure survivorship rules can address edge cases effectively. For example, regulatory compliance may require that Tax ID adheres to legal standards and authoritative sources, while cross-system dependencies may involve attributes like Bank Account Details reliant on external integrations, such as financial institutions. Additionally, some attributes, such as Legal Name, may require manual overrides during conflict resolution. Document these considerations in the Comments column of the Survivorship Rules Template to ensure they are accounted for without compromising data quality.

RACI Matrix – Define Attribute Survivorship Rules

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Steward	Metadata Specialist	Data Owner
4.1.3.A	Identify Key Attributes and Sources	I	A	R	R	C	C
4.1.3.B	Define Survivorship (Retention) Criteria	I	A	R	R		C
4.1.3.C	Define Approaches to Address Complex Conflicts and Exceptions	I	A	R	C		
4.1.3.D	Define Validation Strategies for Survivorship Outcomes	I	A	R	C		
4.1.3.E	Specify Roles for Future Implementation	A	R	I			
4.1.3.F	Document Special Considerations		A	R	C		

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Establish Attribute Survivorship Rules

4.2 MRD Architecture & Modeling

This section will guide the Entity through key activities related to MRD architecture and modeling, including selecting the appropriate architecture archetype (e.g., Registry, Consolidation, Coexistence, or Centralized) based on the defined MRDM requirements. It will also guide the entity in designing and documenting the conceptual and information architecture and gathering technical requirements for the MDM Hub. The process ensures that the chosen architecture not only addresses the current data management challenges but is also scalable to support future business and technological needs.

4.2.1 Select Master and Reference Data Hub Architecture

This section outlines the steps to evaluate and select the appropriate Master and Reference Data Hub architecture to effectively manage master and reference data. The chosen architecture will also define how information flows to downstream or consuming systems, enabling seamless data sharing and operational efficiency.

Process – Select Master and Reference Data Hub Architecture

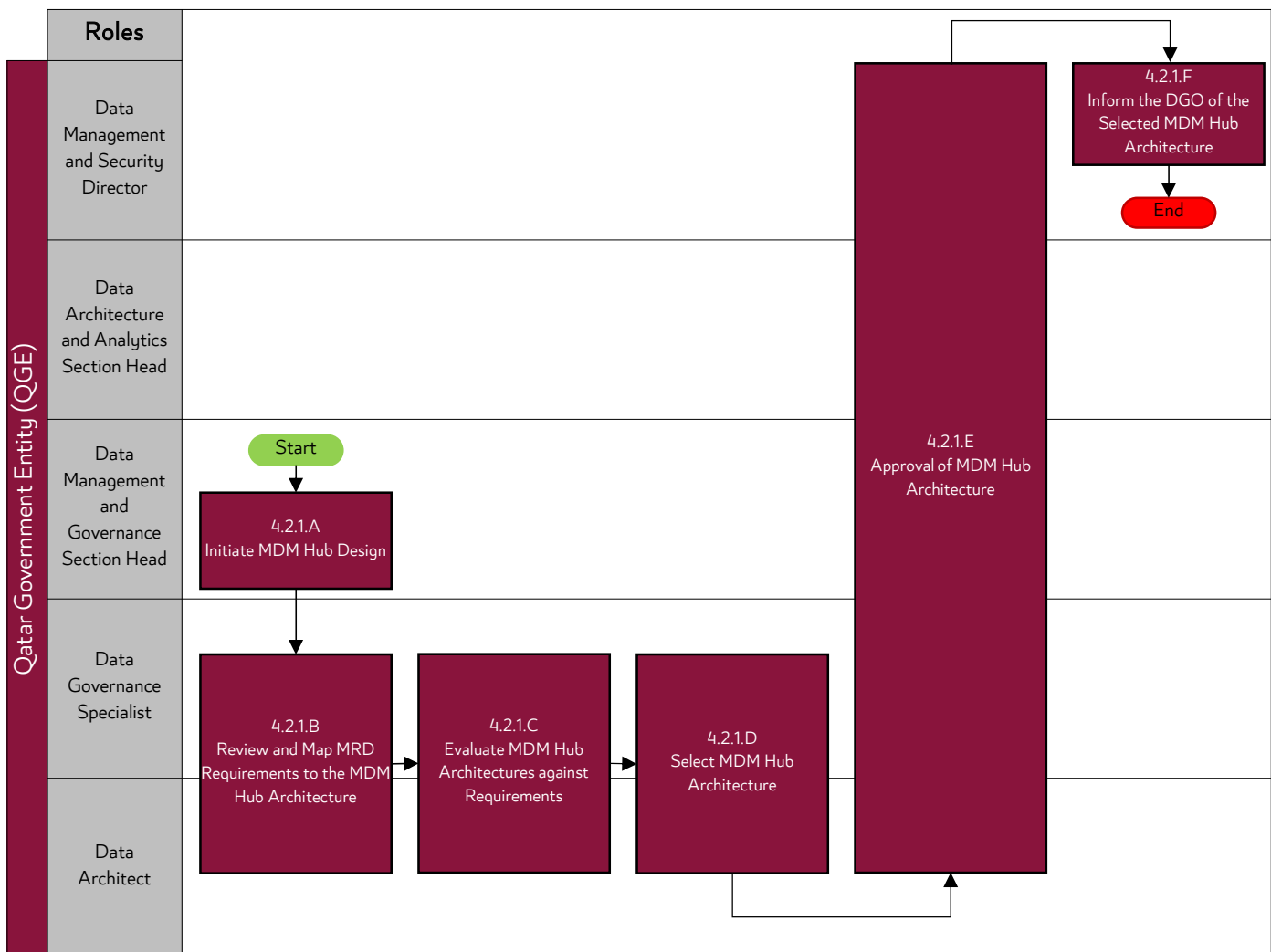


Figure 9: Process Map for Selecting Master and Reference Data Hub Architecture

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

4.2.1.A

Initiate MDM Hub Design: The Data Management and Governance Section Head initiates the MDM hub design exercise and assigns the Data Governance Specialist to oversee and lead the activity.

4.2.1.B

Review and Map MRDM Requirements to the MDM Hub Architecture: The Data Governance Specialist and Data Architect evaluate the identified MRDM requirements (section 3.2.1) and map them against each of the following MDM hub architecture types:

i. Registry Hub

Key Features:

- Master data records remain in their respective source systems.
- The hub stores an index of master data with key identifiers only.
- Creation and modification of master data are maintained in source systems, while the hub provides links to those records.
- The hub includes rules to transform and assemble data dynamically.

When to use the Registry Hub Architecture?

- When minimal disruption to existing systems is required.
- To assemble a real-time single view of master and reference data from various source systems for operational or analytical purposes.

When not to use the Registry Hub Architecture?

- If guaranteed performance and reliability are required for accessing a pre-assembled, single view of master data.
- If a physical MDM database is needed for read/write operations or as a golden record repository.

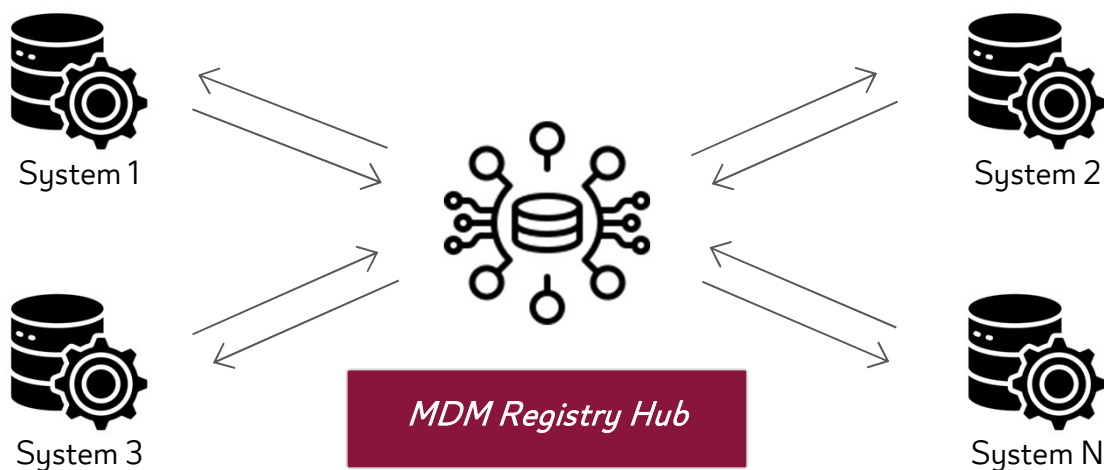


Figure 10: Registry Hub Design

4.2.1.B
Cont.

ii. Consolidation Hub

Key Features:

- Data from various source systems is aggregated into a central MDM hub.
- The MDM system cleanses and deduplicates this data to create a "golden record".
- The "golden record" serves as a read-only reference, with no updates sent back to the source systems.

When to use the Consolidation Hub Architecture?

- When a single authoritative source of data is needed for reporting or analytics.
- To centrally cleanse and deduplicate data for improved quality and consistency.
- When source systems do not require updates from the MDM hub.

When not to use the Consolidation Hub Architecture?

- When master data needs to be harmonized across source several systems by pulling updates from the MDM hub.
- If real-time operational read/write capabilities are required.
- When creating a golden record in a central hub is restricted due to legal or political reasons.

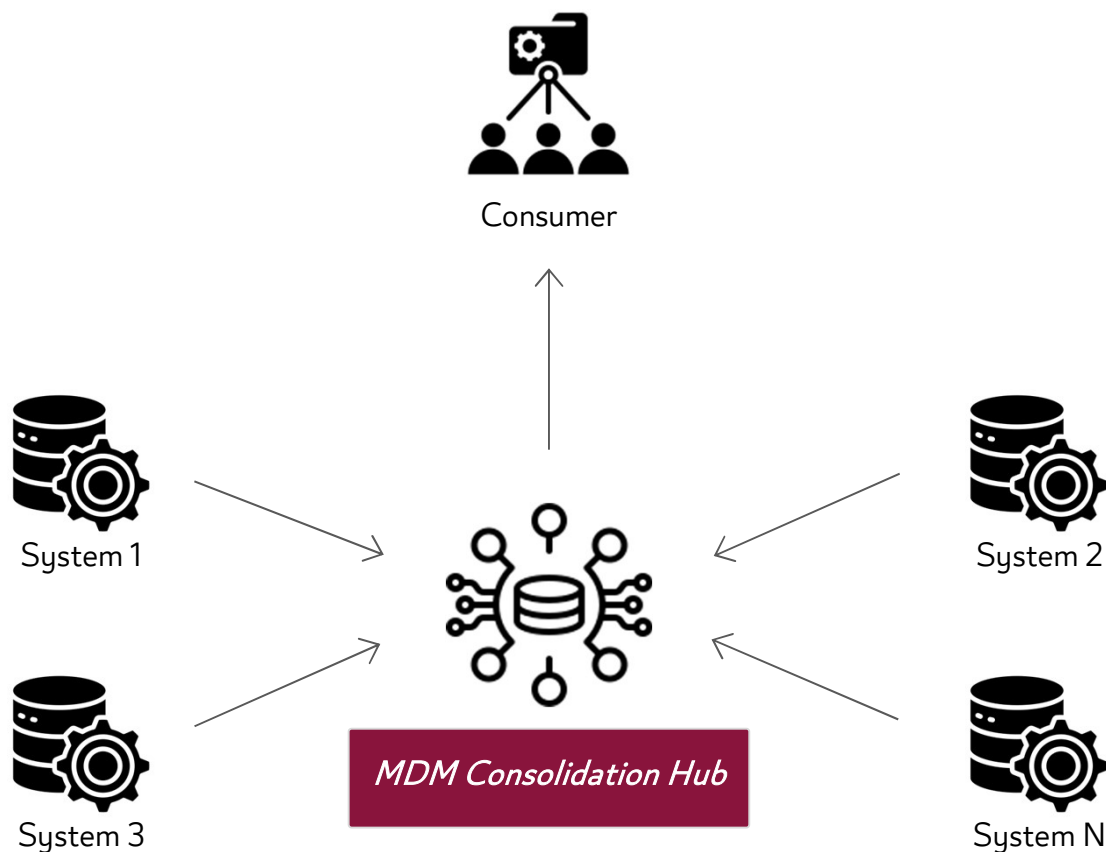


Figure 11: Consolidation Hub Design

4.2.1.B
Cont.

iii. Coexistence Hub

Key Features:

- Two-way synchronization between the hub and source systems ensures both maintain replicas of master data records.
- Creation and modification of master data are maintained in source systems.

When to use the Coexistence Hub Architecture?

- If a physical MDM database is needed for read-only purposes in an operational environment.
- To improve and manage the quality (completeness, consistency) of master data across multiple systems.
- When the IT infrastructure supports middleware for required workflows.

When not to use the Coexistence Hub Architecture?

- If the MDM system must directly support transaction-processing requirements.
- If regulatory or operational constraints restrict the creation of a golden record in a central MDM hub.

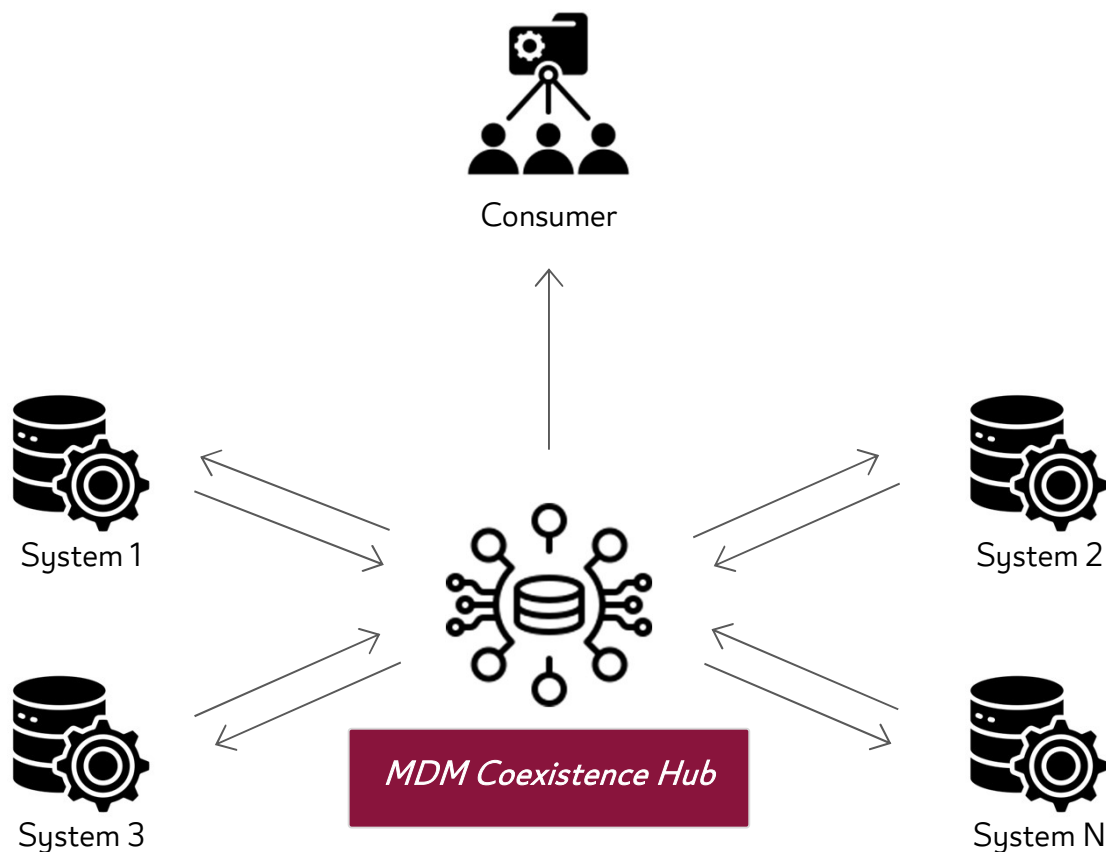


Figure 12: Coexistence Hub Design

4.2.1.B
Cont.

iv. Centralized Hub

Key Features:

- Master data records are fully migrated to the hub, which acts as the single source of truth.
- Creation, modification, and consumption of master data occur exclusively in the hub.
- All consumer applications rely on the hub for their master data needs.

When to use the Centralized Hub Architecture?

- If an MDM system is required to support transaction-processing operations.
- When senior-level sponsorship is available to ensure buy-in for enterprise-wide governance.
- When resources and stakeholder support exist for implementing a centralized hub.

When not to use the Centralized Hub Architecture?

- If there is insufficient stakeholder support, senior-level buy-in, or funding.
- If legal or regulatory constraints restrict the creation of a centralized golden record.

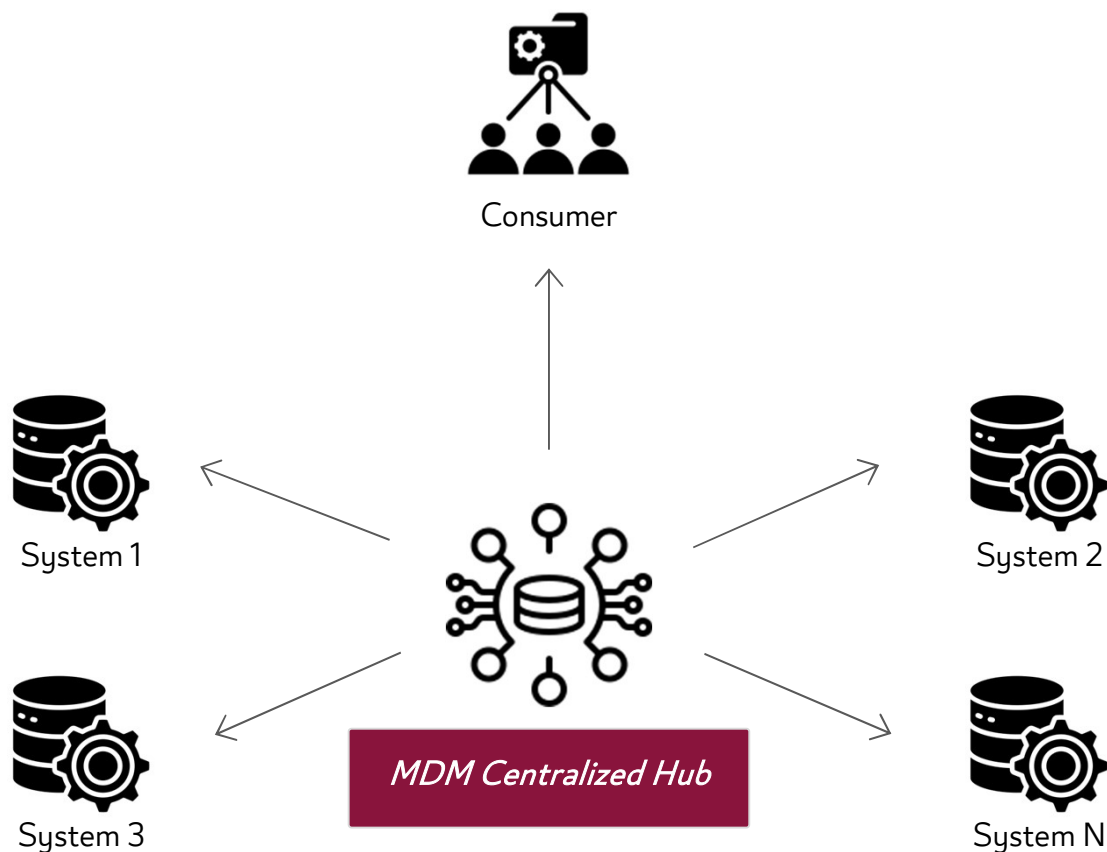


Figure 13: Centralized Hub Design

4.2.1.B
Cont.

Map the Hub Architecture MRDM Requirements (filter “Hub Architecture” within the “Requirement Type” column) from section 3.2.1 to each of the MDM hub architecture types by populating the **MRDM Design Decision Template Sheet** in the **MRD Requirements Template** to identify the most suitable option.

Asset Guidance

The **MRD Requirements Template** includes an **MRDM Design Decision Template Sheet**, which the Entity can use to map the MRDM requirements to the various MDM Hub architecture types.

Illustrative

MRDM Design Decision					
The following table can be used to map the MRD Requirements to each of the MRD hub architectures					
Field	Description				
Requirement ID	Unique number to distinctly identify the requirement.				
MRD Requirements	Detailed description of the strategic requirements to be considered when selecting the most suitable MRD hub design.				
Registry Hub	Assessment of whether the Registry Hub architecture can fulfill the MRD requirement (Yes, Partially, No).				
Consolidation Hub	Assessment of whether the Consolidation Hub architecture can fulfill the MRD requirement (Yes, Partially, No).				
Coexistence Hub	Assessment of whether the Coexistence Hub architecture can fulfill the MRD requirement (Yes, Partially, No).				
Centralized Hub	Assessment of whether the Centralized Hub architecture can fulfill the MRD requirement (Yes, Partially, No).				

Requirement ID	MRD Requirement	Registry Hub	Consolidation Hub	Coexistence Hub	Centralized Hub



Key Message / Information

Mapping MRD requirements to these hub architectures ensures the selected design aligns with the Entity's functional, technical, and governance needs. Consider the Entity's current IT infrastructure, data governance priorities, and operational objectives when selecting the most suitable architecture.

4.2.1.C

Evaluate MDM Hub Architectures against Requirements: Each hub architecture is evaluated against the Entity's defined set of requirements to determine the most suitable option for the Entity. Beyond the defined requirements, other key evaluation factors may include:

- **Integration Requirements:** Ability to integrate with upstream and downstream systems in real-time or batch modes.
- **Scalability:** Support for future data growth and new data domains.
- **Implementation Feasibility:** Alignment with the Entity's existing IT systems, operational complexity, and resource capacity.
- **Governance Alignment:** Compliance with national data standards, enterprise data model, and security requirements.

4.2.1.D

Select MDM Hub Architecture: Based on the evaluation, the Data Governance Specialist and stakeholders select the MDM hub architecture that best meets the Entity's requirements. The selected architecture should:

- Address the challenges, requirements, and priorities identified in 3.2.1.
- Provide a balance between operational efficiency and governance needs.
- Align with the Entity's strategic goals and national data management standards.

4.2.1.E

Approval of MDM Hub Architecture: The selected MDM Hub Architecture design is presented to the Data Management and Security Director, Data Management and Governance Section Head, and Data Architecture and Analytics Section Head for approval. This review ensures the technical feasibility of the design and verifies that the MDM Hub Architecture will effectively meet the specified MRDM requirements.

4.2.1.F

Inform the DGO of the Selected MDM Hub Architecture: Provide the DGO with a summary of the chosen MDM Hub Architecture, including the rationale behind its selection.

RACI Matrix – Select Master and Reference Data Hub Architecture

Indicative

Sr No.	Activity	DGO	Data Management and Security Director	Data Management and Governance Section Head	Data Architecture and Analytics Section Head	Data Governance Specialist	Data Architect
4.2.1.A	Initiate MDM Hub Design	I	A	R	I	I	I
4.2.1.B	Review and Map MRD Requirements to the MDM Hub Architecture			A		R	R
4.2.1.C	Evaluate MDM Hub Architectures against Requirements			A		R	R
4.2.1.D	Select MDM Hub Architecture			A		R	R
4.2.1.E	Approval of MDM Hub Architecture		A/R	R	R	R	R
4.2.1.F	Inform the DGO of the Selected MDM Hub Architecture	I	A/R				

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Select MDM Hub Design

4.2.2 Develop MDM Conceptual Architecture

The development of the MRD Conceptual Architecture ensures that all identified MRDM requirements are addressed through a high-level architectural design. The process includes iterative collaboration with stakeholders to refine the architecture for alignment with Entity's goals, data standards, and technical feasibility.

Process – Develop MDM Conceptual Architecture

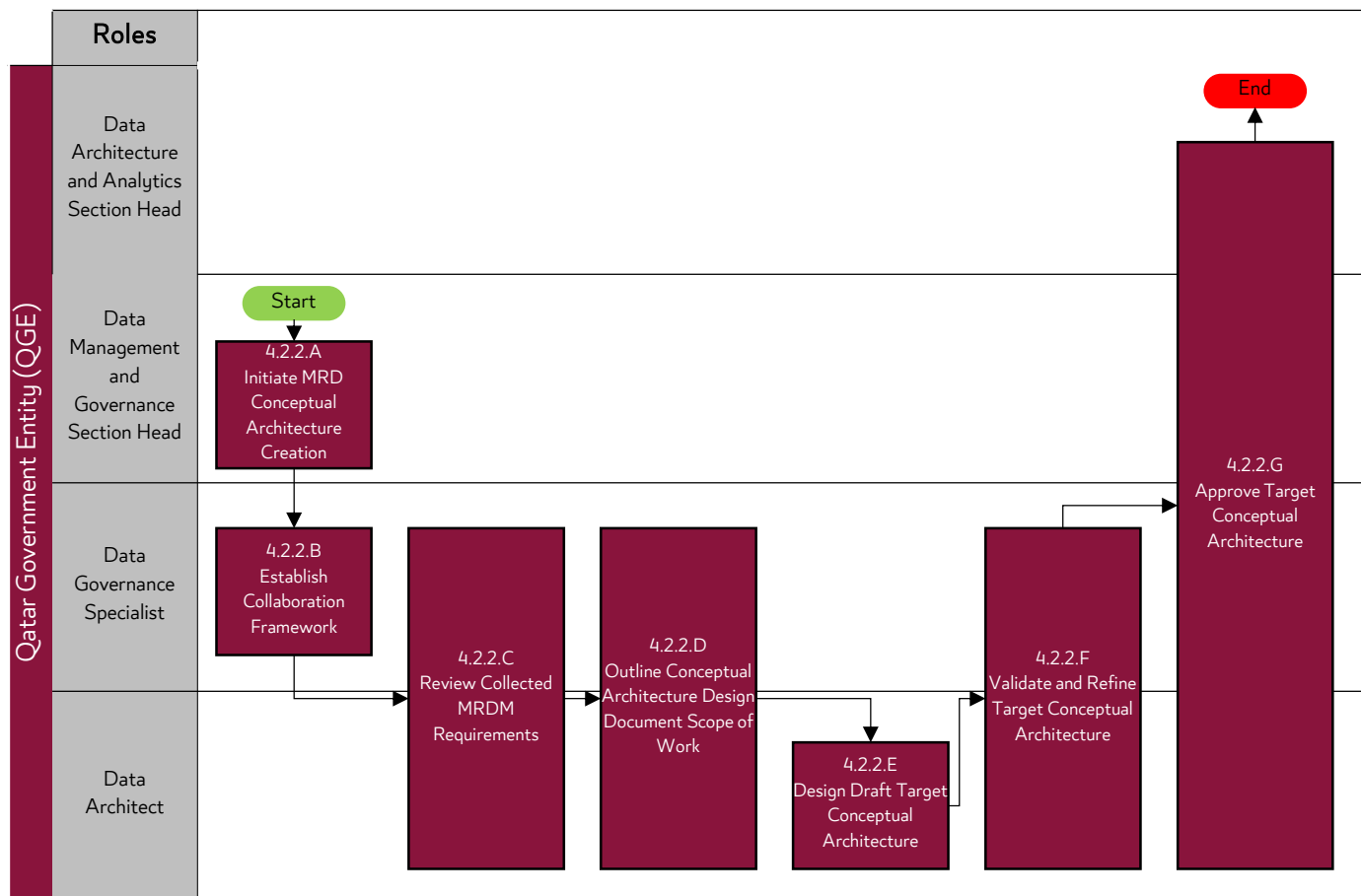


Figure 14: Process Map for Developing MDM Conceptual Architecture

Legend: ■ Process Step ◆ Decision ● Process Start ● Process End □ Document ■ Off-Page Reference

4.2.2.A

Initiate MDM Conceptual Architecture Creation: The Data Management and Governance Section Head initiates the development of the MDM conceptual architecture and assigns the Data Governance Specialist and the Data Architect to oversee and execute the activity.

4.2.2.B

Establish Collaboration Framework: Identify key stakeholders (e.g., Data Stewards, System Owners, and IT Leads) and schedule a kick-off meeting to define the process for architectural review (e.g., timelines, feedback mechanisms).

4.2.2.C

Review Collected MRDM Requirements: The Data Governance Specialist and Data Architect reviews and analyzes the collected MRDM requirements to ensure that they are comprehensive, categorized, and prioritized. Areas requiring clarification or additional input need to be highlighted and addressed before proceeding.

4.2.2.D

Outline Conceptual Architecture Design Document Scope of Work: The Data Governance Specialist and Data Architect define the scope of work for the conceptual architecture design document, including:

- i. **Architecture Description:** Description of the overall defined architecture concept.
- ii. **Architecture Component Definitions and Descriptions:** Building blocks (the MDM Hub, data quality engine, consuming applications, etc.) of the MDM conceptual architecture with a description of each component purpose.
- iii. **Conceptual Architecture Diagram:** High-level view of how components work together to address the Entity's MRDM requirements.

4.2.2.E
(1/3)

Design Draft Target Conceptual Architecture: The Data Architect collaborates with Data Stewards and System Owners to develop the Target Conceptual Architecture.

Provide a concise overview of the conceptual architecture, highlighting how it meets the Entity's needs. Explain its purpose, functionalities, and design rationale, ensuring stakeholders understand its alignment with MRDM objectives, governance policies, and technical workflows. In alignment with the defined scope:

- i. **Develop Architecture Description:**
 - a. **Clarify Architecture Concept:**
 - Describe what the architecture aims to achieve, its scope, and the data flows it will manage.
 - Provides stakeholders with a business and technical rationale for the architecture.
 - b. **Highlight Key Functionalities:**
 - Summarize the core features (e.g., data ingestion, data quality management, hierarchy management).
 - Links these functionalities to organizational goals, such as improved data quality or governance.
- ii. **Define Architecture Components (sample components include, but are not limited to):**
 - **Data Sources:** Systems or repositories that generate, store, or maintain raw data (e.g., ERP, CRM, third-party data feeds). These provide the inputs for MRD processing.

4.2.2.E
Cont.
(2/3)

- **Data Ingestion Layer:** Mechanisms to acquire and import data into the MDM Hub (MRD staging area), including ETL (Extract, Transform, Load) processes, APIs, and streaming data pipelines.
- **Data Quality Engine:** A system or module responsible for profiling, validating, and standardizing data before entering the MRD staging area, where the match and merge rules are executed.
- **MDM Staging Area:** Temporary storage within the MDM Hub where ingested data is processed, transformed, and validated before it is consolidated into the MRD repository.
- **Match and Merge Engine:** A module responsible for identifying duplicate records, applying survivorship rules, and consolidating data to create and maintain golden records for MRD.
- **Hierarchy Manager:** A component of the MDM Hub that enables the creation, visualization, and management of relationships between master data entities (e.g., parent-child structures, organizational hierarchies, or product groupings).
- **Workflow Engine:** A system that manages workflows for tasks such as issue resolution, data approval, exception handling, and escalation, ensuring stewardship processes are executed effectively.
- **Metadata Repository:** A centralized store for business and technical metadata, enabling data lineage, traceability, and clear understanding of the data assets within the MDM Hub.
- **MRD Repository:** The core component of the MDM Hub that stores cleansed, validated, and consolidated MRD, serving as the "single source of truth" for consuming systems.
- **Data Stewardship Integration:** Tools and interfaces that enable data stewards to define rules, monitor data quality, resolve issues, and enforce governance policies across the MRD ecosystem.
- **Integration Layer:** Middleware, APIs, and / or data pipelines that enable the synchronization of MRD with downstream consuming systems (e.g., analytics platforms, operational systems).
- **Data Security and Access Controls:** Security mechanisms to govern access to data within the MDM Hub, ensuring that only authorized users can view or modify MRD

iii. Develop Conceptual Architecture Diagram:

- Create a visual representation that includes:
 - **Interactions Between Components:** Highlight how each architectural layer integrates (e.g., Data Quality Engine connected to the Workflow Engine).
 - **Data Flow:** Show data movement across ingestion, processing, storage, and distribution layers.
 - **System Interdependencies:** Identify upstream and downstream systems interacting with the MDM hub.

4.2.2.E
Cont.
(3/3)

Key Message / Information

Reference the Data Architecture & Modeling National Data Standards to ensure compliance (e.g., reference architecture, naming conventions, interoperability protocols).



Asset Guidance

The **MRD Requirements Template** contains the **Conceptual Architecture Sheet** which the Entity can leverage to provide a concise overview of the conceptual Architecture.

Illustrative

MRD Conceptual Architecture	
The following table can be used to provide a concise overview of the conceptual architecture.	

Field	Description
Component ID	Unique number to distinctly identify the conceptual architecture component.
Conceptual Architecture Components	Name of the conceptual architecture component
Description	Detailed description of the conceptual architecture components
Dependency/ Flow of information	Other conceptual components of the architecture it interacts

Component ID	Conceptual Architecture Components	Description	Dependency/ Flow of information

4.2.2.F

Validate and Refine Target Conceptual Architecture: The Data Governance Specialist and Data Architect lead the review of the target conceptual architecture to ensure that all identified requirements are addressed and fulfilled by the proposed design. This includes:

i. Review Sessions with Data Stewards

- Conduct validation workshops with Data Stewards to ensure that the architecture addresses operational requirements, governance needs, and data quality expectations.
- Capture feedback regarding key MRD functionalities (e.g., data quality validation rules, match/merge configurations).

ii. Technical Review with System Owners

- Collaborate with IT system owners to validate the feasibility of integrating proposed architecture components with existing infrastructure.
- Address any potential performance bottlenecks or integration challenges.

iii. Iterative Refinement

- Update the architecture based on stakeholder feedback and validation findings.
- Document decisions/justifications for inclusion/exclusion of specific functionalities.

4.2.2.G

Approve Target Conceptual Architecture: The target conceptual architecture is presented to the Data Management and Governance Section Head and Data Architecture and Analytics Section Head for approval. The Data Management and Governance Section Head evaluates whether the architecture meets the identified MRD requirements, while the Data Architecture and Analytics Section Head assesses the technical feasibility of the architecture and its alignment with the national data architecture and modeling standards.

RACI Matrix – Develop MDM Conceptual Architecture

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Architecture and Analytics Section Head	Data Governance Specialist	Data Architect	Data Stewards	System Owners
4.2.2.A	Initiate MRD Conceptual Architecture Creation	A	R	I	I	I	I	I
4.2.2.B	Establish Collaboration Framework		A		R	I	I	I
4.2.2.C	Review Collected MRDM Requirements		A	I	R	R	I	I
4.2.2.D	Outline Conceptual Architecture Design Document Scope of Work		A	I	R	R		
4.2.2.E	Design Draft Target Conceptual Architecture		A			R	C	C
4.2.2.F	Validate and Refine Target Conceptual Architecture		A		R	R	C	C
4.2.2.G	Approve Target Conceptual Architecture	A	R	R	R	R		

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Develop Conceptual Architecture

4.2.3 Develop MDM Information Architecture

This section outlines the steps to define and document the Information Architecture for the Entity's MDM Hub, aligning with its selected hub architecture. It ensures accurate ingestion, transformation, validation, and distribution of MRD while maintaining high data quality and governance standards. This architecture forms the foundation for efficient data management, ensuring consistency and accessibility across the enterprise.

Process – Develop MDM Information Architecture

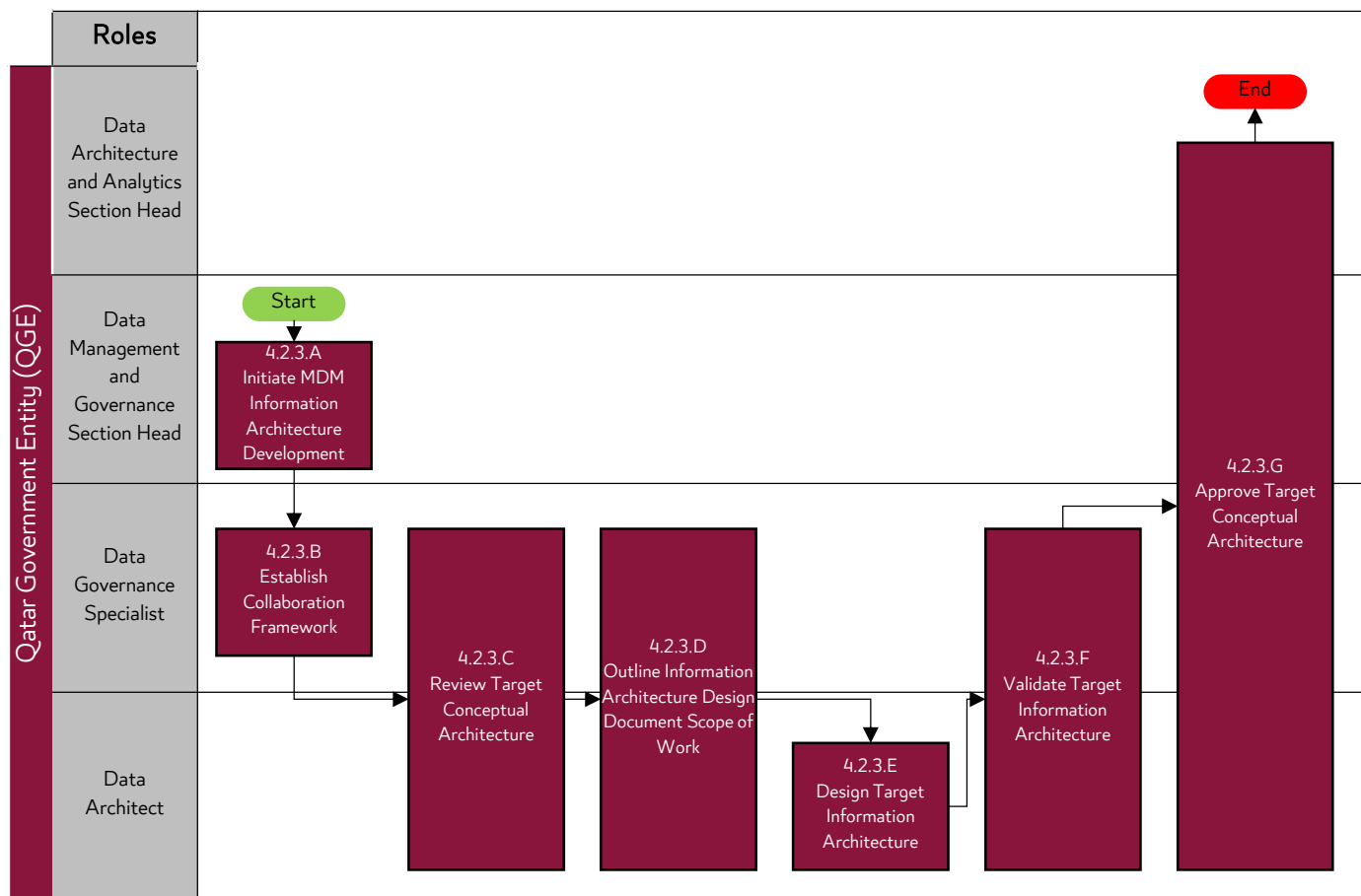


Figure 15: Process Map for Developing MDM Information Architecture

Legend: Process Step (Maroon Rectangle), Decision (Maroon Diamond), Process Start (Green Oval), Process End (Red Oval), Document (White Rectangle with Maroon Border), Off-Page Reference (Maroon Rectangle with White Border)

4.2.3.A

Initiate MDM Information Architecture Development: The Data Management and Governance Section Head initiates the development of the MDM information architecture, designating the Data Architect and Data Governance Specialist to lead and oversee its execution.

4.2.3.B

Establish Collaboration Framework: Engage key stakeholders and initiate discussions to align on the objectives and scope of the MDM information architecture. Establish structured collaboration mechanisms for collecting technical inputs, addressing conflicts, and reviewing architectural deliverables. Clearly define stakeholder responsibilities and timelines to ensure effective contributions throughout the process.

4.2.3.C

Review Target Conceptual Architecture: The Data Architect reviews the approved conceptual architecture, verifying all components and data flows. If any technical gaps or missing details emerge, they are identified and addressed to ensure a smooth transition to the information architecture phase.

4.2.3.D

Outline Information Architecture Design Document Scope of Work: The Data Governance Specialist and Data Architect define the scope of work for the Information Architecture Design Document, ensuring alignment with the approved conceptual architecture. This document provides a detailed framework for structuring, managing, and integrating MRD within the Entity, covering key areas including but not limited to:

- i. **Architecture Description:** Defines the purpose, vision, and alignment of the MDM information architecture with business, governance, and technical objectives.
- ii. **MRD Domains:** Outlines the required MRD domains to be brought into the MDM hub, including attributes, classification, and governance metadata.
- iii. **Data Sources and Ingestion:** Outlines the source systems, integration mechanisms, and ingestion processes for bringing data into the MDM Hub.
- iv. **Data Quality Management:** Describes the Data Quality Engine, which performs validation, cleansing, and enforcement of DQ rules on source data before it enters the MRD Staging Zone. Only data that meets the defined quality thresholds moves into the MRD environment, ensuring the integrity and reliability of ingested data.
- v. **MRD Landing (Staging) Area:** Details the temporary storage area where validated data is prepared for processing. This includes staging rules, schema validation, and initial transformation steps before integration into the core MRD repository.
- vi. **Data Modeling and Transformation:** Defines the MRD domains along with their attributes, classifications, and governance metadata. Specifies the conceptual and logical data models, including entities, relationships, and transformation rules.
- vii. **Golden Record and Hierarchy Management:** Establishes the matching, merging, and survivorship rules for creating Golden Records, ensuring that duplicate records are resolved, and data is harmonized to create a single authoritative version. Additionally, defines the hierarchical structures governing parent-child relationships across master data entities (e.g., customer → account → transaction), and describes how these hierarchies are managed and updated within the MDM Hub.
- viii. **Data Stewardship User Interface:** describes the role-based access controls, governance workflows, and interface functionalities that enable data stewards to manage and maintain MRD effectively. Outlines the dashboarding, search, update, and approval functionalities available to end users.

4.2.3.D
Cont.ix. **Data Flows: Overview of data flows, including:**

- Master data ingestion from source systems to the MDM Hub.
- Golden Record synchronization between MDM Hub and consuming applications.
- Master and Reference data distribution to downstream applications.

4.2.3.E
(1/3)

Design Target Information Architecture: The Data Architect collaborates with Data Stewards and System Owners to develop the Target Information Architecture.

Provide an overview of the information architecture, highlighting how it meets the Entity's needs. Explain its purpose, functionalities, and design rationale, ensuring stakeholders understand its alignment with MRDM objectives, governance policies, and technical workflows. In alignment with the defined scope:

i. **Develop Architecture Description:**

- Define the overall vision of the MDM Information Architecture. This includes how it will structure and manage MRD across the Entity.
- Provides stakeholders with a business and technical rationale for the information architecture.

ii. **MRD Domains:**

- Document all the required MRD domains (e.g., customer, employee, vendor) to be brought into the MDM hub. For each MRD domain, specify:
 - **Attributes:** Include attribute name, data type, source, and validation rules.
 - **Classification:** Define whether the attribute is operational, analytical, or strategic.
 - **Governance Metadata:** Include ownership, data steward, and lifecycle information.
- Document business rules for data creation, update, and retirement.

**Key Message / Information**

Reference the MRD Inventory defined in the Data Catalog to identify the required MRD objects, datasets, and attributes to be considered in the MDM information architecture. Reference the MRDM Rules Template for the respective details on MRD validation rules.

iii. **Data Sources and Ingestion:**

- **Data Source Inventory:** List all the systems expected to (e.g., ERP, CRM, external APIs) contribute data to the MDM Hub.
- **Integration Mechanisms:**
 - **Batch:** Define batch processing schedules and data formats.
 - **Real-time:** Specify API requirements, authentication, and real-time triggers (if any).
- **Landing Area Design:** Detail how raw data is staged temporarily in the MDM Hub without modifications, enabling validation and profiling.

4.2.3.E
Cont.
(2/3)iv. **Data Quality Management:**

- **Data Quality Rules:** Specify data quality rules to validate incoming data (e.g., email format, mandatory fields).
- **Cleansing Processes:** Describe how data inconsistencies (e.g., duplicates, null values) are resolved.
- **Monitoring and Profiling:** Detail mechanisms for ongoing data profiling and quality checks.



Key Message / Information

Refer to the Data Quality Rules established in the Data Quality Management domain to support this activity. Additionally, consult the Data Quality Management National Data Standards to ensure alignment with established standards and best practices for maintaining data quality and compliance.

iv. **MRD Landing (Staging):**

- **Staging Rules:**
 - **Schema Validation:** Define schema rules to ensure data adheres to expected data formats, types, and structures. Invalid data is flagged for correction or exclusion.
 - **Normalization:** Outline details and processes to normalize data (e.g., if one system records citizen age in years and another in months, convert both to years for consistent data).
 - **Standardization:** Outline details and processes to convert data values into a consistent format (e.g., if one system records “NPC” and another records “National Planning Council” for the same entity, standardization would convert both to “National Planning Council”).

v. **Data Modeling and Transformation:**

- **Conceptual Data Model:**
 - Define high-level entities and relationships (e.g., Customer -> Order -> Payment).
- **Logical Data Model:**
 - Specify detailed relationships, attributes, primary keys, foreign keys, and constraints.



Key Message / Information

Leverage the Data Architecture & Modeling domain guidelines to ensure that data models are compliant with established design principles, systematically defined, organized, and effectively managed to promote consistency, scalability, and alignment with organizational objectives.

4.2.3.E
Cont.
(3/3)

iv. Golden Record and Hierarchy Management:

➤ Golden Record Creation (Match and Merge Engine):

- **Match Rules:** Define and document algorithms, methods, and rules used for deduplication and similarity detection (e.g., exact matching, fuzzy matching)
- **Merge Rules:** Specify survivorship rules and attribute trust levels for merging records and creating the authoritative golden record.

➤ Hierarchy Management:

- **Hierarchy Structures:** Define parent-child relationships for entities (e.g., for products, parent-child relationships may include categories and subcategories).
- **Configuration Rules:** Specify how hierarchies will be maintained and updated within the MDM Hub.



Key Message / Information

Leverage the MRD Rules Template defined in activity 4.1 to support the activity VI (Match Engine and Golden Record Creation).

vii. Data Stewardship User Interface:

- **Role-Based Access Controls:** Define user roles (e.g., Data Steward, Analyst) and their permissions.
- **Functionalities:** List key functionalities for the user interface, including:
 - Search, update, and approve records.
 - Manual matching and merging.
 - Dashboarding for issue tracking and data quality metrics.

ix. Data Flows:

- **Master Data Ingestion:** Specify how data flows from source systems to the MDM Hub, detailing:
 - Trigger mechanisms for real-time and batch updates.
 - Error-handling processes for failed ingestion.
- **Golden Record Synchronization:** Define how Golden Records are updated in consuming systems, including frequency (e.g., real-time, daily).
- **Reference Data Distribution:** Detail how standardized reference data is shared with downstream applications to ensure consistency.

4.2.3.F

Validate Target Information Architecture: The Data Governance Specialist and Data Architect lead the review of the target information architecture to ensure that all identified requirements are addressed and fulfilled by the proposed design.

i. **Review Sessions with Data Stewards:**

- Conduct validation workshops with Data Stewards to confirm that the architecture addresses operational requirements, data governance needs, and workflows for stewardship tasks.
- Capture feedback on key components such as data ingestion processes, data quality validation rules and workflows, match and merge configuration for creating golden records, and user interface elements supporting data steward activities.

ii. **Technical Review with System Owners:**

- Collaborate with System Owners and IT Leads to validate the feasibility of integrating the proposed architecture with the Entity's existing IT infrastructure.
- Address potential challenges related to performance bottlenecks (e.g., system scalability, data throughput), compatibility with upstream and downstream systems, and implementation of data flow mechanisms (e.g., batch processing, APIs).

iii. **Iterative Refinement:**

- Update designs to resolve gaps or inconsistencies identified by stakeholders.
- Incorporate adjustments to workflows, data flows, and component interactions as required.
- Document decisions and justifications for any modifications, ensuring transparency in the refinement process.

4.2.3.G

Approve Target Conceptual Architecture: The target information architecture is presented to the Data Management and Governance Section Head and Data Architecture and Analytics Section Head for approval. The Data Management and Governance Section Head evaluates whether the architecture meets the identified MRD requirements, while the Data Architecture and Analytics Section Head assesses the technical feasibility of the architecture and its alignment with the national data architecture and modeling standards.

RACI Matrix – Develop MDM Information Architecture

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Architecture and Analytics Section Head	Data Governance Specialist	Data Architect	Data Stewards	System Owners
4.2.3.A	Initiate MDM Information Architecture Development	A	R	I	I	I	I	I
4.2.3.B	Establish Collaboration Framework		A		R	I	I	I
4.2.3.C	Review Target Conceptual Architecture		A	I	R	R	I	I
4.2.3.D	Outline Information Architecture Design Document Scope of Work		A	I	R	R		
4.2.3.E	Design Target Information Architecture		A			R	C	C
4.2.3.F	Validate Target Information Architecture		A		R	R	C	C
4.2.3.G	Approve Target Conceptual Architecture	A	R	R	R	R		

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Develop MDM Information Architecture

4.2.4 Capture MDM Hub Technical Requirements

This section outlines the steps to define and document the technical requirements for the Entity's MDM Hub, ensuring alignment with the selected target MDM hub architecture.

Process – Capture MDM Hub Technical Requirements

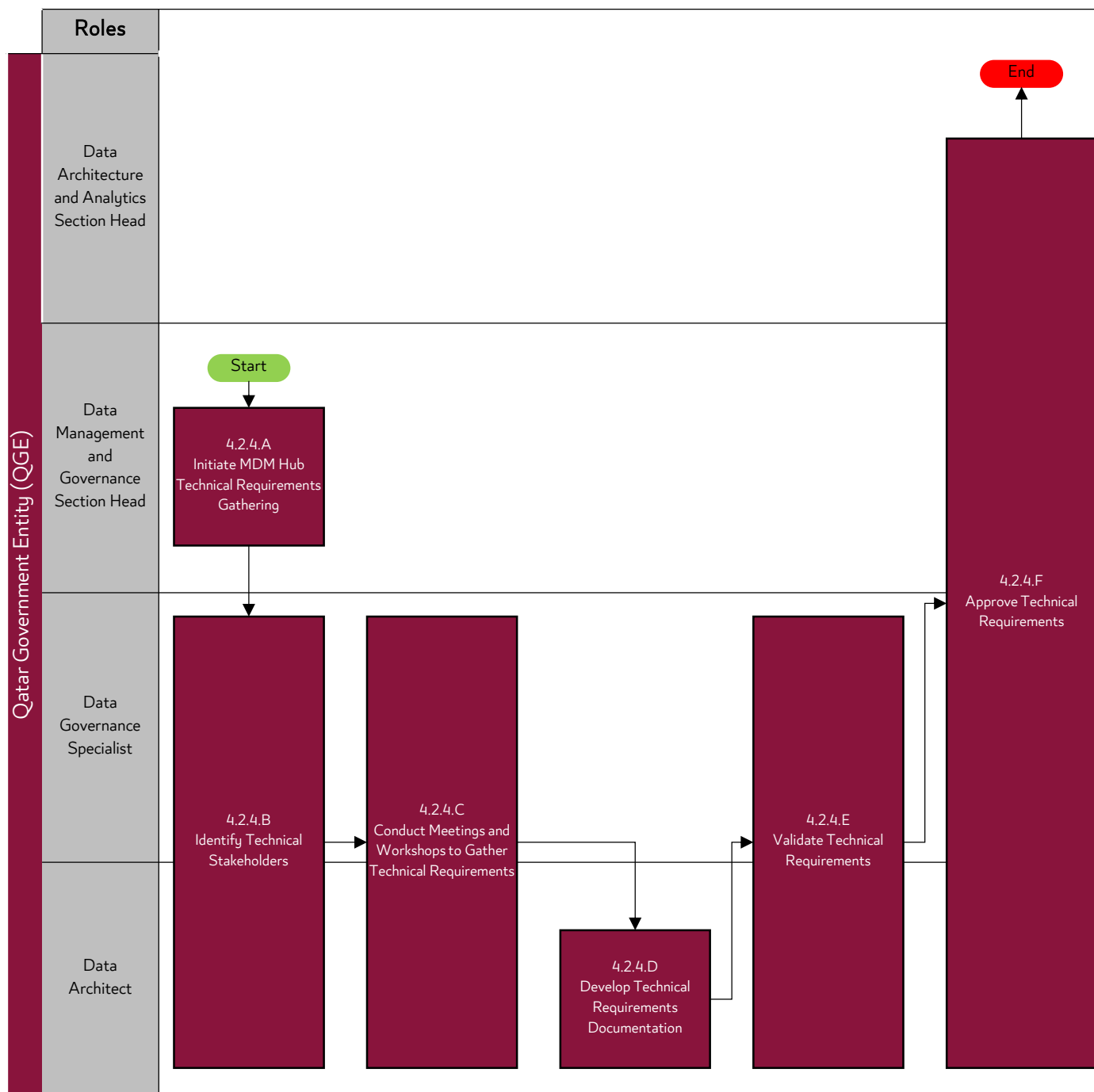


Figure 16: Process Map for Capturing MDM Hub Technical Requirements

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

4.2.4.A

Initiate MDM Hub Technical Requirements Gathering: The Data Management and Governance Section Head assigns the Data Governance Specialist to oversee the technical requirements gathering process. Additionally, the Data Architect is appointed to drive the discussions and ensure alignment with the conceptual and information architecture. The exercise is scoped to focus on the technical capabilities necessary to implement the selected MDM hub architecture.

4.2.4.B

Identify Technical Stakeholders: Identify and engage stakeholders responsible for providing technical input. Key stakeholders include:

- **System Owners:** Represent systems where master and reference data originates, is stored, or consumed.
- **Technical Data Stewards:** Ensure data infrastructure supports technical processes.
- **Solution Architect:** Translate functional requirements into technical specifications.
- **Business Data Stewards:** Provide business context that influences technical requirements.

4.2.4.C

Conduct Meetings and Workshops to Gather Technical Requirements: Facilitate collaborative workshops and discussions with stakeholders to gather technical and functional requirements for the MDM hub. Discussions should address the following areas:

i. **Management of Workflows:**

- Define the approach for creating, modifying, and governing MRD records.
- Clarify the roles and responsibilities of Data Stewards in managing data lifecycles.

ii. **Version Control:**

- Determine the approach for tracking changes to MRD records over time.
- Ensure alignment with both business requirements and the Entity's existing technology stack.

iii. **Security Requirements:**

- Identify security measures to protect data at rest and in motion.
- Discuss requirements for secure data exchanges between the hub and connected systems or applications.

iv. **Functional Capabilities:**

- Define the operational features required from the MDM hub such as mechanisms for importing and exporting MRD, and automation of tasks such as data collection, cleansing, and validation.

v. **Technical Capabilities:**

- Outline the technical features the MDM hub must support. Examples include:
 - **Batch Integration:** Requirements for batch processing, specifying frequency and formats.
 - **Real-Time Integration:** Requirements for real-time data exchanges using APIs.
 - **Near Real-Time Integration:** Requirements for near real-time data integration using middleware (e.g., service bus).

4.2.4.D

Develop Technical Requirements Documentation: The Data Architect compiles the gathered input into a detailed technical requirements document. The document should include:

- **Technical Requirement Descriptions:** Clearly define each requirement, including its purpose and alignment with the MRDM requirements.
- **Functional/Non-Functional Area:** Indicate whether the requirement pertains to creating a specific functionality (functional) or a broader non-functional requirement.
- **Technical Requirement:** Provide a detailed description of the technical requirement, including specifics on what the solution must address.
- **Requirement Criticality:** Prioritize which requirements will be addressed in the implementation.

Asset Guidance

The **MRD Requirements Template** contains the **MRD Technical Requirements Sheet** which the Entity can leverage to provide a concise overview of the conceptual Architecture.

Illustrative

MDM Technical Requirements				
The following table can be used to map detail the MDM Technical Requirements and their criticality				
Field	Description			
Technical Requirement ID	Unique identifier for each technical requirement			
Functional/Non-Functional Area	Indicates whether the requirement pertains to creating a specific functionality (Functional) or a broader, non-functional requirement (e.g., performance, scalability).			
Technical Requirement	Detailed description of the technical requirement, including specifics on what the solution must address.			
Requirement Criticality	Classify the importance of the requirement into categories such as: Must Have, Good to Have, or Future Need. This helps prioritize which			
Requestor	The individual or team that submitted the requirement or identified the need.			
Technical Requirement ID	Functional/Non-Functional Area	Technical Requirement	Requirement Criticality	Requestor

4.2.4.E

Validate Technical Requirements: The Data Governance Specialist and Data Architect lead the review sessions to validate the documented technical requirements:

- Review with Data Stewards:**
 - Ensure alignment with operational needs, governance objectives, and stewardship workflows.
 - Capture feedback on requirements related to data validation, version control, and automation.
- Review with System Owners and Solution Architect:**
 - Validate the feasibility of technical requirements within the Entity's existing technology stack.
 - Address potential challenges related to integration, scalability, and security.

4.2.2.F

Approve Technical Requirements: The finalized technical requirements document is presented to the Data Management and Governance Section Head and Data Architecture and Analytics Section Head for approval. The Data Management and Governance Section Head ensures requirements address business needs and support governance objectives, while the Data Architecture and Analytics Section Head assesses the technical feasibility of the requirements.

RACI Matrix – Capture MDM Hub Technical Requirements

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Architecture and Analytics Section Head	Data Governance Specialist	Data Architect	Data Stewards	System Owners
4.2.4.A	Initiate MDM Hub Technical Requirements Gathering	A	R	I	I	I	I	I
4.2.4.B	Identify Technical Stakeholders		A		R	R	I	I
4.2.4.C	Conduct Meetings and Workshops to Gather Technical Requirements		A	I	R	R	I	I
4.2.4.D	Develop Technical Requirements Documentation		A	I	I	R		
4.2.4.E	Validate Technical Requirements		A		R	R	C	C
4.2.4.F	Approve Technical Requirements	A	R	R	R	R		

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Capture MDM Hub Technical Requirements

5. Master and Reference Data Management Implementation



5. Master and Reference Data Management Implementation

What is Master and Reference Data Management Implementation and why does the Entity need it?

MRDM implementation involves selecting an MRDM tool and building the MDM Hub to govern, integrate, and ensure data quality. This includes installing and configuring the tool, setting up access controls, defining data models, applying validation rules, and integrating with enterprise systems. The MDM Hub centralizes data consolidation and synchronization, ensuring interoperability. Rigorous testing, validation, and training ensure smooth adoption and compliance with National Data Standards.

As the Entity matures its data management capabilities, these tools can be enhanced to include advanced features like API integration, real-time data synchronization, dataset version mapping, and support for hierarchical datasets and multilingual data. For entities seeking cutting-edge capabilities, the tools can be further expanded to incorporate AI/ML-powered functionalities, such as automated discovery and entity resolution.

By implementing these tools, the Entity establishes a robust framework for managing master and reference data across its systems, ensuring data accuracy, consistency, and accessibility. These tools provide a scalable solution that evolves with the Entity's needs, enabling better decision-making, operational efficiency, and compliance with the National Data Standards.

The need for MRDM Implementation arises from several critical factors:

1. **Data Consistency:** Ensures a single source of truth and uniformity for master and reference data across systems, reducing inconsistencies and enabling reliable data usage across the Entity
2. **Data Quality:** Implements validation, matching, and cleansing processes to eliminate duplicates, standardize records, and improve the overall quality of master and reference data
3. **Integration and Interoperability:** Facilitates seamless data exchange between systems through APIs, batch processing, and real-time synchronization, ensuring smooth interactions between source and consuming applications
4. **Operational Efficiency:** Streamlines processes for managing master and reference data by automating repetitive tasks such as data entry, validation, and matching, reducing manual effort and improving efficiency

Implementing Master and Reference Data Management Tools enables the Entity to maintain high-quality, standardized data across all systems, supporting consistent and reliable decision-making.

The remaining sections of this chapter will detail the below processes that will guide the Entity on executing the below:

5.1 Master & Reference Data Management Tools

5.1 Master & Reference Data Management Tools

The objective of this process is to guide the Entity in selecting and implementing a master and reference data management tool and building the MDM Hub.

5.1.1 Master & Reference Data Management Tool Selection and Implementation Plan

The objective of this process is to guide the Entity in selecting the most suitable MDM tool.

Process – Master & Reference Data Management Tool Selection and Implementation Plan

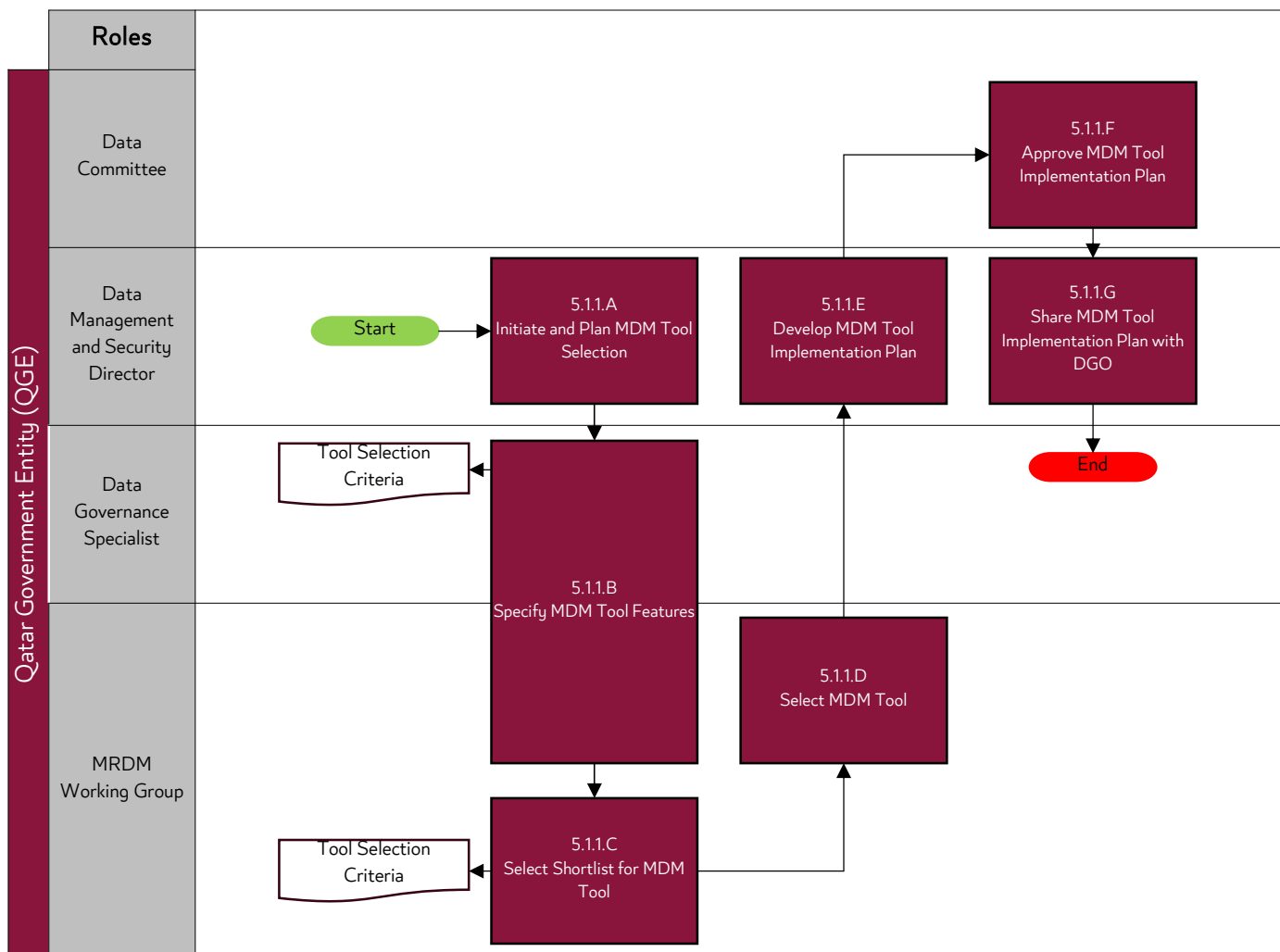


Figure 17 – Master & Reference Data Management Tool Selection and Implementation Plan Process Map

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

5.1.1.A

Initiate and Plan MDM Tool Selection: Define the objectives for the MDM Tool and establish a timeline with key milestones for its selection and implementation. Ensure alignment with the Entity's data strategy and requirements as per the National Data Standards. Form an MRDM Working Group comprising members from data governance, IT, compliance, and business analysis, along with data stewards. Engage this group throughout the selection and implementation process to provide expertise, ensure compliance, and support informed decision-making.

5.1.1.B

Specify MDM Tool Features: Align the MDM tool requirements with the MDM hub technical requirements and architecture to ensure they meet the foundational, advanced, and very advanced features specified in the National Data Standards, which are the following:

i. **Foundational Features:**

- **Workflow Management:** Manage and streamline workflows related to master and reference data.
- **Version Control:** Support for maintaining multiple versions of master and reference data to track changes over time.
- **Data Import and Export:** Facilitate import and export of master and reference data, supporting formats like XML.
- **Validation and Matching:** Validate and match data during entry or batch processing to ensure data accuracy and eliminate duplicates.
- **Web Services Integration:** Enable integration with external systems using web services for data exchange and interaction.

ii. **Advanced Features:**

- **API Integration:** Facilitate integration with external systems using APIs for seamless data exchange.
- **Dynamic Data Synchronization:** Support real-time synchronization of data across distributed systems to maintain consistency.
- **Dataset Version Mapping:** Provide capabilities to map and manage transitions between dataset versions.
- **Hierarchical Dataset Support:** Enable management of data organized in hierarchical structures to address complex relationships.
- **Multilingual Data Handling:** Support reference data in multiple languages to meet diverse operational needs.
- **Customizable User Interface:** Offer customizable interfaces to cater to specific user and operational requirements.

iii. **Very Advanced Features:**

- **AI-Powered Master and Reference Data Management:** Leverage AI/ML capabilities for automated discovery and identification of master data, including pre-trained models for entity resolution and advanced matching techniques.

Once requirements are identified, populate the MDM Tool Selection Criteria Template, specifying the tool features, feature weightings, and cost weightings to guide the selection process.

5.1.1.B
Cont.

 Asset Guidance

The **MDM Tool Selection Criteria Template** helps the Entity document mandatory and additional MDM tool features, list various tool options, specify feature and cost weightings based on importance, and guide the Entity through the decision-making process.

Illustrative

	Master and Reference Data Management Tool Selection Criteria Template
The following tables provides a structured template to guide the evaluation and selection of a master and reference data management tool	

Field	Description
Tool feature	The specific tool feature required based on the domain's standards (e.g., version control).
Tool feature weight	The weighted percentage assigned to the tool feature based on importance to the respective entity.
Feature level of Advancement	The level of advancement of the feature as per the standards (Foundational, Advanced, or Very Advanced).
[Tool name 1,2,3,X]	Whether the suggested tool has the feature described (Yes/No).

Features Scoring	Description
1 - Low	Limited functionality, insufficient to meet key requirements
2 - Medium	Basic functionality, meets most requirements but with some limitations
3 - High	Strong functionality, meets all requirements effectively with minor gaps
4 - Very High	Comprehensive functionality, meets all requirements with advanced features and flexibility

Pricing Scoring	Description
1 - Low	High cost solution
2 - Medium	Moderate cost solution
3 - High	Budget friendly solution
4 - Very High	Low cost solution

Tool Feature	Tool Feature Weight	Feature level of Advancement	[Tool name 1]	[Tool name 2]	[Tool name 3]	[Tool name X]
Cost	30%	Foundational	4			
Workflow management	30%	Foundational	3			
Version control	10%	Foundational	2			
Master and reference data import and export capabilities (e.g., XML)	10%	Foundational	2			
Validation and matching at data entry	10%	Foundational	1			
Web services integration	10%	Foundational				
Total:			65%	0%	0%	0%

5.1.1.C

Select Shortlist for MDM Tool: Review and evaluate available MDM tool options to create a shortlist that meets the defined requirements. Populate the **MDM Tool Selection Criteria Template** from the previous step with the identified tool names and assign scores for each tool based on the specified features.

5.1.1.D

Select MDM Tool: Assess the shortlisted tools based on feature and pricing scoring, then finalize the selection of the MDM tool that best aligns with the Entity's data strategy and National Data Standards requirements.

5.1.1.E ***Develop MDM Tool Implementation Plan:*** Create a comprehensive implementation plan for the MDM tool, detailing the phases and budget for its installation, implementation, and configuration, as well as the development, validation, and testing of the MDM Hub built on the tool. The plan should also cover user training on the MDM Hub and outline clear timelines, responsibilities, and resource requirements to ensure a smooth transition to the new system.

5.1.1.F ***Approve MDM Tool Implementation Plan:*** Present the MDM tool's implementation plan and budget for review and approval. Ensure alignment with organizational priorities, budget constraints, and National Data Standard requirements.

5.1.1.G ***Share MDM Tool Implementation Plan with DGO:*** Submit the finalized MDM Tool Implementation Plan to the National Data Governance Office (DGO) to keep the DGO informed about the Entity's planned implementation timeline and key milestones and ensure compliance to the National Data Standards.

RACI Matrix – Master & Reference Data Management Tool Selection and Implementation Plan *Indicative*

Sr No.	Activity	DGO	Data Management and Security Director	Data Management and Governance Section Head	MRDM Working Group	Data Governance Specialist	Data Committee
5.1.1.A	Initiate and Plan MDM Tool Selection		A/R	I		C	C
5.1.1.B	Specify MDM Tool Features		I	I	R	A/R	
5.1.1.C	Select Shortlist for MDM Tool		I	A	R	I	
5.1.1.D	Select MDM Tool		C	A	R	C	
5.1.1.E	Develop MDM Tool Implementation Plan		A/R	C	C	C	
5.1.1.F	Approve MDM Tool Implementation Plan		C	I	I	C	A/R
5.1.1.G	Share MDM Tool Implementation Plan with DGO	I	A/R				

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Specify MDM Tool Features and Select Tool



Develop MDM Tool Implementation Plan

5.1.2 Implement and Establish MDM Hub on Selected MDM Tool

This section outlines the steps required to implement the MDM Hub on the selected MDM tool. To ensure successful implementation, the Entity must assign a project manager to oversee and track the progress of the MDM Hub implementation, consult with the Data Architect, and establish a MRD Working Group comprising MDM developers, and MDM testers. This working group can be onboarded on a temporary basis to drive the project. Additionally, select employees within the Entity should be trained to provide post-implementation support and manage the MRD program in alignment with the MRD lifecycle process.

Process – Implement and Establish MDM Hub on Selected MDM Tool

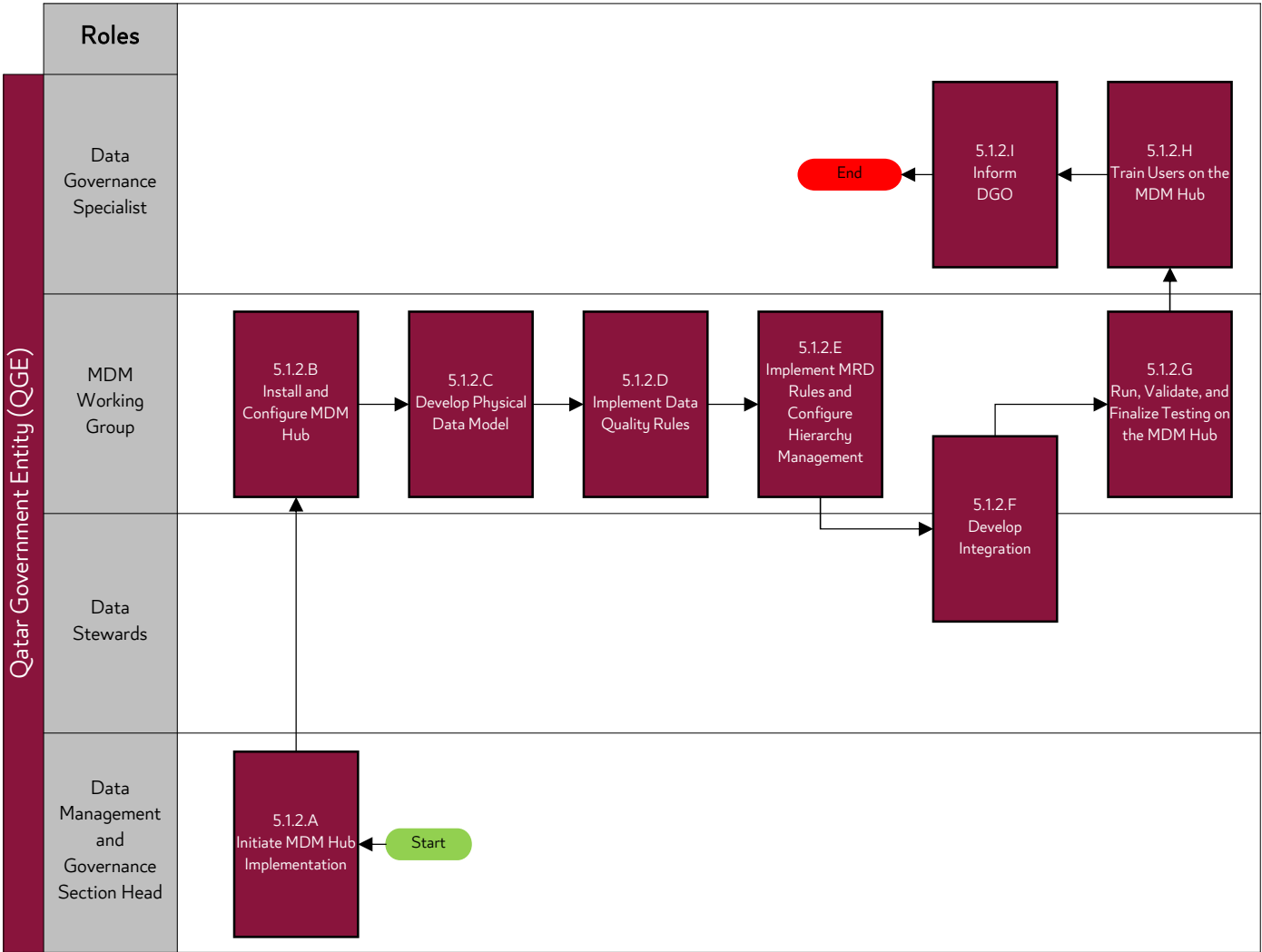


Figure 18 – Process Map for Implementing and Establishing the MDM Hub on the Selected MDM Tool

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

5.1.2.A

Initiate MDM Hub Implementation: The Data Management and Governance Section Head initiates the implementation of the MDM Hub by appointing a Project Manager to oversee the program. The Section Head may delegate responsibility to the Data Governance Specialist to effectively manage the implementation process and ensure alignment with governance and technical requirements.

5.1.2.B

Configure MDM Hub: The MDM Developers proceed with installing the MDM tool and configuring the MDM Hub to ensure it is fully operational and meets governance and security standards. This includes::

- i. **Install the MDM Tool:** Deploy the selected MDM platform and verify successful installation.
- ii. **Define Users and Roles:** Establish user groups and role-based permissions, ensuring access is aligned with governance policies.
- iii. **Configure Role-Based Access Management:** Assign access levels based on user responsibilities, ensuring secure and controlled interactions with the MDM Hub.
- iv. **Configure Error Handling:** Implement mechanisms to detect, log, and manage data processing errors and exceptions.
- v. **Configure Audit Logging:** Enable tracking of changes, updates, and data modifications to maintain transparency and compliance.
- vi. **Configure MDM Workflows and Alerts:** Set up automated workflows for match and merge processes, data quality validation, and approvals with real-time alerts for key stakeholders.
- vii. **Configure Backup and Archival Processes:** Develop backup and archival scripts to ensure data integrity, recovery readiness, and compliance with retention policies.

5.1.2.C

Develop Physical Data Model: The data modelers and architects create the physical Master Data Model within the MDM tool, following the specifications defined in the MDM Information Architecture. This activity is carried out in consultation with the Data Architect to ensure continuous alignment with the national data standards.

5.1.2.D

Implement Data Quality Rules: The MDM Developers configure Data Quality (DQ) Rules in the MDM Hub to ensure incoming data meets predefined standards, in alignment with the National Data Quality Management Standards.

5.1.2.E

Implement MRD Rules and Configure Hierarchy Management: The MDM Developers implement the MRD rules previously defined, ensuring their correct application within the MDM Hub. These rules govern how data is matched and merged to create the golden records. Additionally, the MDM Developers configure Hierarchy Management to establish structured relationships between master data entities. This includes:

- i. **Establish Parent-Child Relationships:** Define and configure hierarchical structures between master data entities (e.g., customer accounts, product categories, organizational units).
- ii. **Maintain Hierarchy Integrity:** Ensure that hierarchical relationships are accurate, logically structured, and aligned with business requirements.
- iii. **Enable Hierarchy Navigation and Reporting:** Configure settings that allow users to traverse and analyze hierarchical relationships efficiently.

5.1.2.F

Develop Integrations: The MDM Developers and Data Stewards establish connectivity between the MDM Hub and other systems / applications, as specified in the information architecture. This includes:

- i. **Developing APIs:** Create APIs for secure data retrieval, updates, and synchronization between systems.
- ii. **Building Data Pipelines:** Implement batch and real-time data ingestion pipelines to integrate source data into the MDM Hub.
- iii. **Connecting Downstream Applications:** Ensure ERP, CRM, analytics platforms, and other consuming applications can efficiently access master data.

5.1.2.G

Run, Validate, and Finalize Testing on the MDM Hub: The MDM Testers execute comprehensive test cases to validate the functionalities of the MDM Hub, ensuring all components perform as expected before deployment. This includes:

- i. **Testing Data Quality Rules:** Verify that validation, cleansing, and enrichment processes are correctly applied.
- ii. **Running Match and Merge Tests:** Ensure duplicate detection, similarity thresholds, and survivorship rules function as expected.
- iii. **Validating Integrations:** Test API connections, data pipelines, and system interoperability.
- iv. **Executing Performance and Security Tests:** Assess system scalability, response times, and role-based access enforcement.

All issues identified during testing are logged, tracked, and managed by the MDM Testers. The MDM Developers address and resolve defects, after which the resolved items are sent back for retesting. This process continues iteratively until all identified issues are remediated.

Once retesting confirms that all defects have been fixed, the MDM Testers provide a formal sign-off and publish the final test report. Following this, the Data Governance Specialist presents the solution for a final review by key stakeholders, including Data Owners and Data Stewards. These stakeholders validate the test results against the initial functionality requirements.

Upon final approval, the implemented solution is cleared for deployment to the production environment for GO-LIVE.

5.1.2.H

Train Users on the MDM Hub: Conduct training sessions for all relevant stakeholders, including data stewards, business users, and IT personnel, to ensure effective use and adoption of the MDM Hub. The training should focus on equipping users with the necessary skills to manage MRD effectively, covering critical functionalities such as data quality checks, match and merge processes, hierarchy management, and data provisioning. Additionally, provide guidance on role-based access controls, workflow management, and real-time synchronization features to maximize the Hub's potential in supporting Entity data governance and operations.

5.1.2.I

Inform DGO: After the solution is deployed and goes live, the Data Governance Specialist informs the National Data Governance Office (DGO). This communication includes a summary of the implementation, key test results, and a confirmation that the solution aligns with the National Data Standards.

RACI Matrix – Implement and Establish MDM Hub on Selected MDM Tool

Indicative

Sr No.	Activity	DGO	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	MRD Working Group	Data Architect	Data Quality Specialist	Data Owners	Data Stewards	IT Team
5.1.2.A	Initiate MDM Hub Implementation		A	R	I	I					
5.1.2.B	Install and Configure MDM Hub			I	A	R					
5.1.2.C	Develop Physical Data Model			I	A	R	C				
5.1.2.D	Implement Data Quality Rules			I	A	R		C			
5.1.2.E	Implement MRD Rules and Configure Hierarchy Management			I	A	R					
5.1.2.F	Develop Integration			I	A	R				R	C
5.1.2.G	Run, Validate, and Finalize Testing on the MDM Hub		C	C	A	R	I		I	I	C
5.1.2.H	Train Users on the MDM Hub		C	A	R	I	I		C	C	C
5.1.2.I	Inform DGO	I	I	I	R						

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Build MDM Hub on MDM Tool



Test MDM Hub before Deployment

6. Master and Reference Data Performance Management



6. Master and Reference Data Performance Management

What is a Master and Reference Data Performance Management and why does the Entity Need it?

Master and Reference Data Performance Management is the process of continuously monitoring, maintaining, and assessing the quality and effectiveness of master and reference data across the Entity. This involves assigning stewardship roles, defining and tracking key performance indicators (KPIs), implementing monitoring and reporting processes, and establishing service-level agreements (SLAs) to ensure data accuracy, consistency, and compliance. By regularly evaluating data quality and cataloging performance, the Entity can identify areas for improvement and drive initiatives to enhance the reliability and usability of master and reference data.

The need for Master and Reference Data Performance Management arises from several critical factors:

1. **Data Quality and Consistency:** Ensures that master and reference data remain accurate, complete, and aligned with the Data Quality Management standards, supporting consistent usage and reducing errors.
2. **Effective Data Stewardship:** Assigns clear roles and responsibilities to data stewards, enabling them to actively monitor data quality, manage metadata, and facilitate cross-departmental collaboration for data accuracy.
3. **Continuous Improvement:** Establishes KPIs and reporting mechanisms to track performance and support ongoing enhancements, aligning with national standards and promoting data-driven improvements.
4. **Accountability through SLAs:** Defines performance expectations with SLAs to uphold data quality standards and ensure timely resolution of issues, fostering accountability and transparency.

Implementing structured Master and Reference Data Performance Management supports better decision-making, improves operational efficiency, and ensures compliance with data governance requirements. It provides a foundation for reliable, high-quality data that can be trusted and leveraged for strategic initiatives across the Entity.

The remaining sections of this chapter will detail the below processes that will guide the Entity on executing the below:



Guiding Principles MRD Performance Management

1. **Align with Data Quality Standards:** Ensure that performance management efforts adhere to established data quality standards, promoting accuracy, completeness, and consistency in master and reference data.
2. **Empower Data Stewardship:** Assign clear roles and responsibilities to data stewards for monitoring and managing master and reference data quality. Effective stewardship ensures accountability and supports continuous improvement initiatives.
3. **Focus on Data Integrity and Accuracy:** Regularly validate and monitor master and reference data to ensure it meets quality thresholds, minimizing errors and inconsistencies that could impact organizational decision-making.
4. **Implement Continuous Monitoring and Reporting:** Establish a structured framework for ongoing monitoring and reporting, using dashboards, scorecards, and alerts to track data quality KPIs and identify areas needing improvement.
5. **Enforce Accountability through SLAs:** Set up service-level agreements (SLAs) that define performance criteria for data quality and outline expectations for issue resolution. SLAs foster accountability and provide clear standards for managing master and reference data.
6. **Plan for Scalability and Future Requirements:** Design performance management practices that can adapt to new data domains, business needs, and technological advancements.

6.1 MRDM Stewardship & Maintenance Process

MRDM Stewardship & Maintenance Process is an important component of Master and Reference Data Management. This process establishes clear roles and responsibilities for data stewards, enabling them to monitor, maintain, and ensure the quality of master and reference data across the Entity. Data stewards are responsible for verifying data accuracy, completeness, and consistency, aligning with the entity's Data Quality Policy.

6.1.A

Assign and Document Roles and Responsibilities for Data Stewardship: Assign dedicated data stewards for master and reference data and document their roles and responsibilities. This includes defining ongoing activities expected of data stewards to support the quality and maintenance of master and reference data. Some key responsibilities include:

- Monitoring and validating the quality and integrity of Master and Reference Data to ensure they are accurate, complete, and up to date in alignment with the Entity's Data Quality standards
- Analyzing records to identify anomalies after the matching-merging process, and providing feedback to source system owners to improve data accuracy and consistency
- Participating in the development and refinement and fine-tuning of algorithms and processes used in Master & Reference Data Management tools
- Facilitating cross-departmental collaboration to establish and maintain common values and definitions for Master & Reference Data
- Managing metadata for Master & Reference Data, including steward name, originating Entity, expected frequency of updates, schedule for updates, and processes using the Master & Reference Data
- Overseeing workflows for the review and approval of changes to Master & Reference Data

6.1.B

Identify Additional Activities and Assign Roles as Needed: Assess the data stewardship process to determine if additional activities are necessary to fully support master and reference data management. For any identified activities, assign specific roles and responsibilities to ensure they are effectively managed.

RACI Matrix – MRDM Stewardship & Maintenance Process

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Owners	Data StewardS	Data Quality Specialist
6.1.A	Assign and Document Roles and Responsibilities for Data Stewardship	I	A	R	C	I	C
6.1.B	Identify Additional Activities and Assign Roles as Needed	I	A	R	C	I	C

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Assign Roles and Responsibilities to Data Stewards

6.2 MRDM KPIs

As part of the Master & Reference Data Management KPIs specification, the entity shall establish key performance indicators (KPIs) to measure the effectiveness of Master & Reference Data Management efforts and guide continuous improvement initiatives. KPIs shall include, at minimum:

- Master and reference data quality metrics, as defined in the data quality domain

6.2.A

Establish Key Performance Indicators (KPIs) for Master & Reference Data Management:

Define and establish KPIs to measure and track the process of master and reference data management, aligning with the data quality standards. These KPIs should include:

- **Data Quality Levels Across Dimensions:** Track quality metrics for accuracy, completeness, consistency, timeliness, validity, and uniqueness in master and reference data.
- **Data Synchronization Rate:** Track how often master and reference data is updated across integrated systems, ensuring consistency across platforms.
- **System Downtime Impact on MDM Processes:** Monitor the frequency and duration of MDM & RDM system unavailability, assessing its effect on business operations.
- **Data Standardization Compliance:** Measure the percentage of master and reference data adhering to defined naming conventions and data format standards, ensuring consistency and governance compliance.
- **MRDM Error Fix Rate:** Evaluate how efficiently data stewards resolve MDM issues (e.g., data conflicts, corrections), providing insights into issue resolution effectiveness and system stability.

Asset Guidance

The **Data Management KPIs Register Template** includes a table that provides a brief overview of identified KPIs. The Entity can use this template to define each KPI, along with its calculation methodology, target value, and status.

Illustrative

 Data Management KPIs Register Template								
The following table is a template containing what is needed to provide a brief overview of the performance of data management initiatives								
Field	Description							
KPI ID	A unique identifier assigned to each Key Performance Indicator (KPI)							
KPI Name	A short, descriptive name that summarizes what the KPI measures							
KPI Type	Operational or Strategic KPI (if applicable)							
Calculation Methodology	The formula / approach (if any) used to calculate the KPI value							
Target Value	The predefined value or benchmark to achieve for the KPI							
Last Updated Date	The date when the KPI's current value was last recorded or updated							
Status	On Track/Off Track							
Comments	Additional notes or explanations regarding the KPI's performance							
Owner	Outlines whether the business or IT is the owner of the KPI							
KPI ID	KPI Name	KPI Type (if applicable)	Calculation Methodology	Target Value	Last Updated Date	Status	Comments	Owner

6.2.A
Cont.**Key Message / Information**

The Entity must establish a clear process for reviewing and updating KPIs for Master and Reference Data Management to ensure they remain relevant and aligned with evolving business needs and data quality standards. This process should include setting regular intervals for KPI evaluation, incorporating stakeholder feedback, and assessing the effectiveness of each KPI in measuring data quality across dimensions such as accuracy, completeness, and consistency. By periodically revisiting these metrics, the Entity can refine its approach to address emerging challenges, ensure compliance with data quality standards, and leverage advancements in data management practices to drive improvements in master and reference data management.

RACI Matrix – MRDM KPIs*Indicative*

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Steward
6.2.A	Establish Key Performance Indicators (KPIs) for Master & Reference Data Management	I	A	R	R

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Establish MRDM KPIs

6.3 MRDM Monitoring & Reporting

This activity will guide the Entity through the steps to establish Master & Reference Data Management Monitoring and Reporting to assess the quality and effectiveness of master and reference data efforts. At the foundational level, the Entity will align data quality monitoring with established standards, using dashboards and scorecards to visualize performance against key metrics. For advanced monitoring, this activity includes generating process control reports and implementing automated alerts to notify stakeholders of quality issues. Together, these practices ensure data accuracy, promote accountability, and support continuous improvement in master and reference data quality and management.

Process – MRDM Monitoring & Reporting

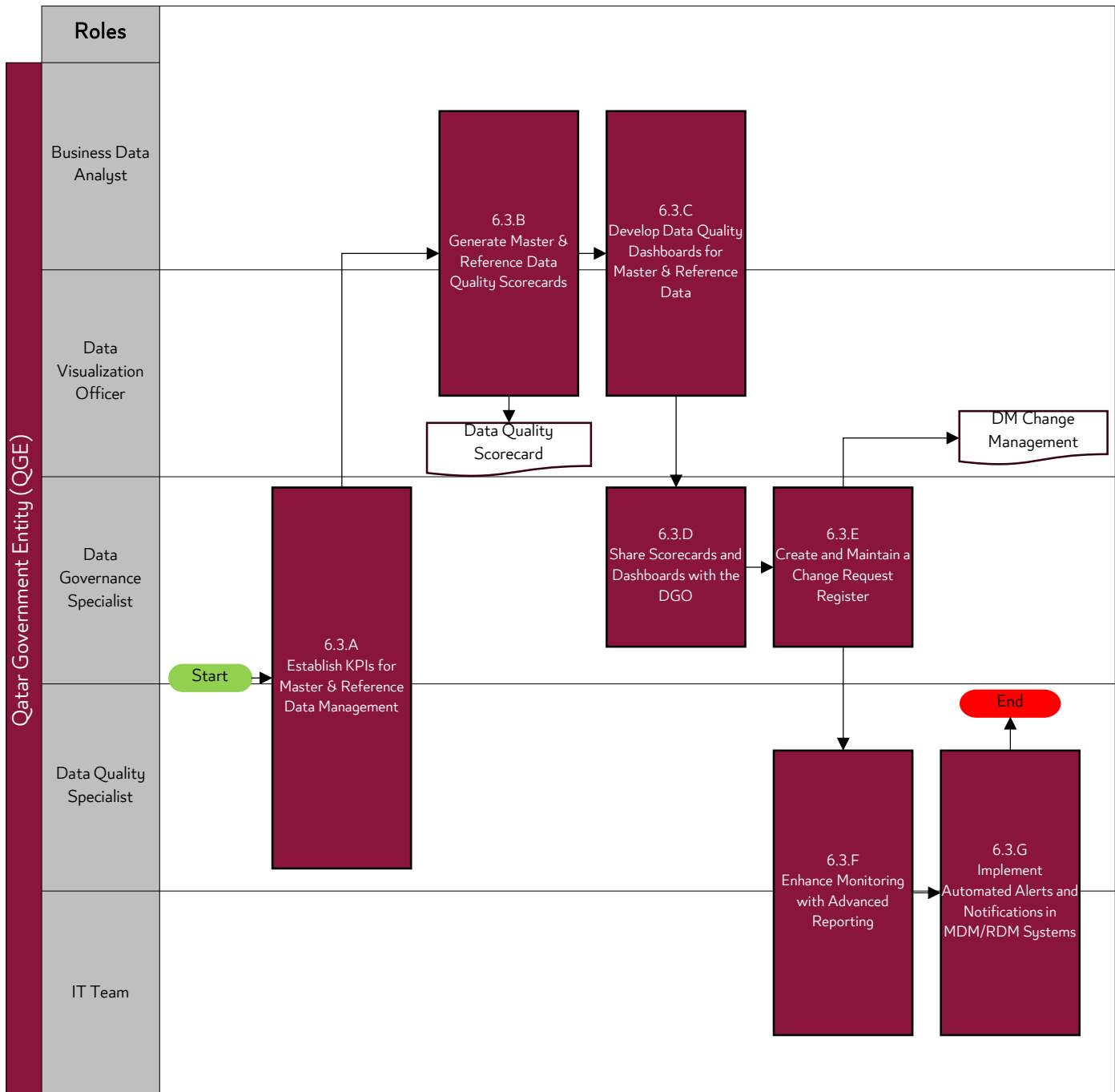


Figure 19: MRDM Monitoring & Reporting Process Map

Legend: Process Step (Maroon rectangle), Decision (Maroon diamond), Process Start (Green oval), Process End (Red oval), Document (White rectangle with black border), Off-Page Reference (Maroon rectangle with black border)

6.3.A

Establish Master & Reference Data Quality Monitoring Aligned with Data Quality Standards:

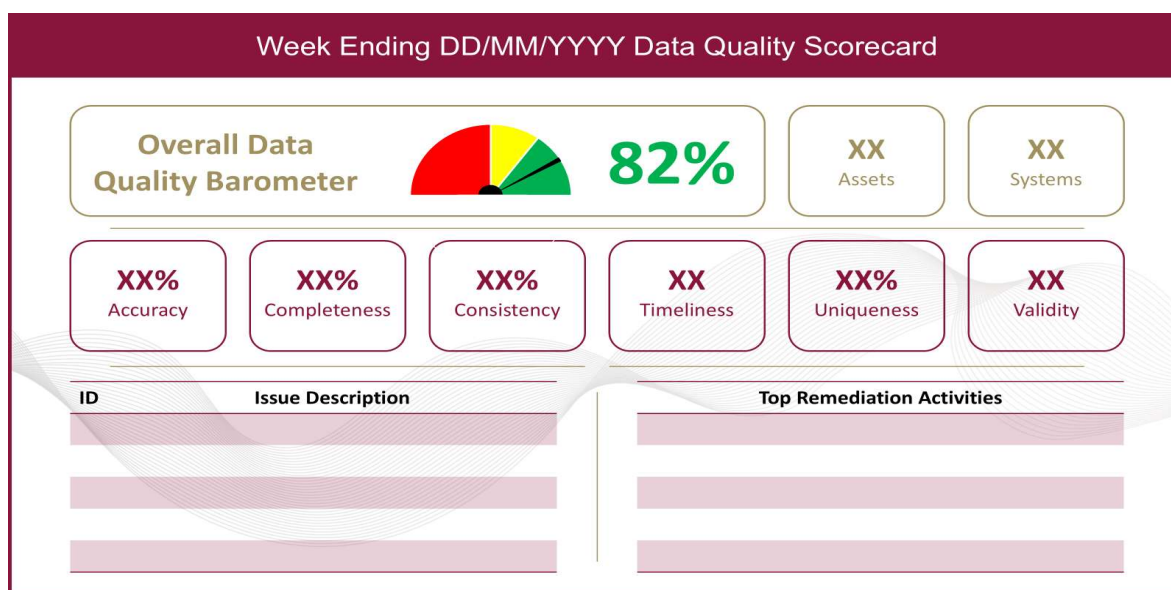
Align master and reference data quality monitoring with the Data Quality Management (DQM) standards, ensuring all monitoring is based on defined data quality rules and KPIs. This includes setting up processes to regularly assess master and reference data against quality thresholds for accuracy, completeness, consistency, timeliness, validity, and uniqueness.

6.3.B

Generate Master & Reference Data Quality Scorecards: Develop and implement a scorecard report that compares master and reference data quality metrics against pre-established thresholds. Use visual indicators (e.g., red-yellow-green) to highlight any data quality issues that do not meet the standards, making it easy to identify and prioritize areas requiring improvement.

 Asset Guidance

The **Data Quality Scorecard Template** is a type of report which focuses on comparing the Data Quality metrics with the thresholds, highlighting areas which do not meet the thresholds (typically with red-yellow-green colors). It also summarizes open data quality issues and concluded remediation activities

Illustrative

6.3.C

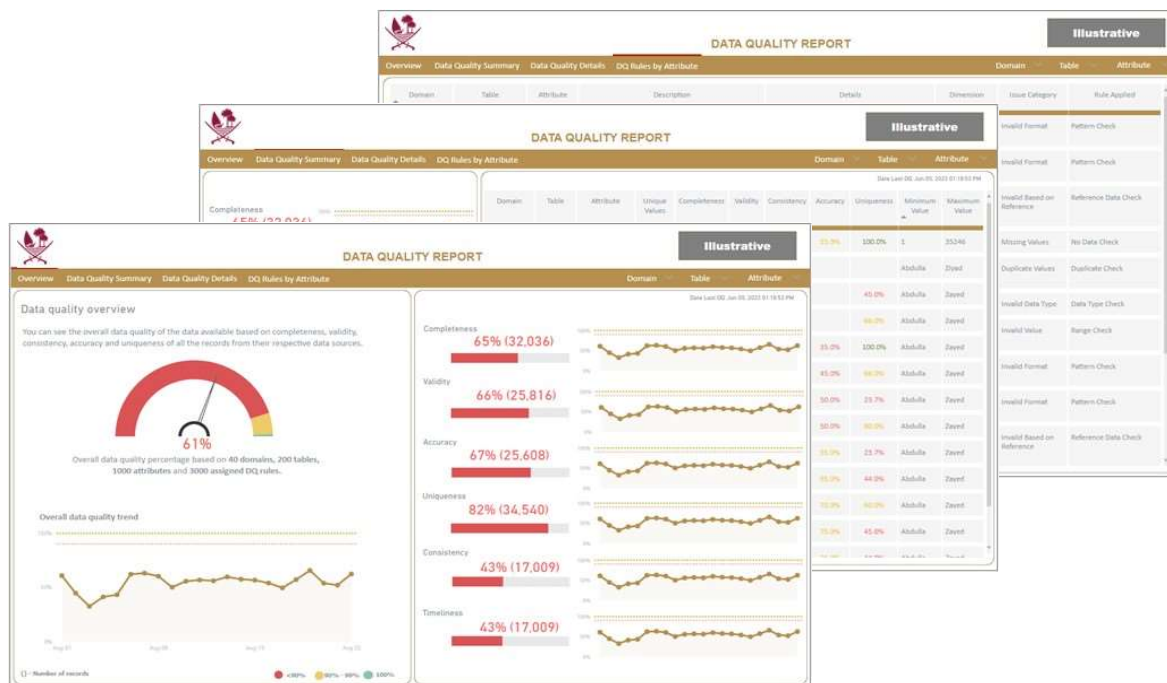
Develop Data Quality Dashboards for Master & Reference Data: Develop an Executive data quality Dashboard for master & reference data with the following characteristics:

- **Executive Summary View:** Create a high-level overview of master and reference data quality metrics, allowing executives to quickly assess the health of critical data assets without needing to delve into operational specifics
- **Aggregated Insights:** Aggregate data from multiple sources, presenting summarized scores for key data quality dimensions like completeness, accuracy, and consistency, providing a holistic view of master and reference data health
- **Visual Clarity:** Use intuitive visualizations such as bar charts and pie charts to convey trends and insights. This approach makes it easier for executives to interpret complex data quality metrics quickly and effectively
- **Key Performance Indicators (KPIs):** Highlight essential KPIs that measure the effectiveness of master and reference data quality initiatives, enabling executives to monitor performance and make informed strategic decisions
- **Trend Analysis:** Include historical comparisons and trend visualizations to show how master and reference data quality has evolved over time, aiding in forecasting, resource allocation, and long-term planning for data quality improvement

Asset Guidance

A **Data Quality Dashboard** provides real-time insights into the quality of data within an entity. It displays key metrics such as accuracy, completeness, and consistency, enabling stakeholders to monitor performance at a glance. The dashboard facilitates quick identification of data quality issues and trends, supporting proactive remediation efforts. By centralizing data quality information, it enhances decision-making and drives continuous improvement initiatives.

Illustrative



6.3.D

Share Scorecards and Dashboards with the DGO: Share the data quality Scorecards and Dashboards with the National Data Governance Office (DGO) on a monthly basis, enabling the DGO to monitor progress, track improvements, and ensure alignment with National Data Standards.

6.3.E

Create and Maintain a Change Request Register: Develop a register to record all change requests for master and reference data, documenting decisions made, actions taken, and their impact. This register provides traceability for modifications, ensuring transparency in decision-making.

Asset Guidance

The **Data Management Change Management Template** contains a Change Request & Impact Assessment Template that the Entity can leverage for documenting changes related to master and reference data management.

Illustrative

Change Request & Impact Assessment					
The following table serves as a template for documenting changes related to data management and their impact					
Field	Description				
Change ID	Unique identifier for the change request or project				
Change Title	A short, descriptive title summarizing the change request or project				
Change Description	Detailed description of the requested change, including what is being modified or introduced and the rationale for the change				
Request Date	The date when the change request was submitted				
Requested By	Name and department of the individual or team requesting the change				
Change Owner	Person or department responsible for managing and overseeing the change implementation				
Priority Level	The urgency of the change request (e.g., High, Medium, Low)				
Impact Level/Change Severity	Estimated impact or severity of the change (High, Medium, Low) on departments, systems, or processes				
Impacted Departments/Teams	List of departments, teams, or individuals impacted by the change				
Change Type	Type of change (Process Change, Technology Change, Role Change)				
Impacted Areas	Specific areas affected by the change (Systems, Processes, Data, Roles)				
Stakeholders Affected	Key stakeholders who will be impacted by the change				
Potential Risks	Risks associated with the change (e.g., Resistance to Change, System Downtime)				
Mitigation Approach	Strategies to mitigate risks (e.g., training, early communication, technical support)				
Requested Implementation Date	Proposed date by which the change should be implemented				
Change Qualification Status	Classification indicating whether the change request meets the necessary criteria for further review and assessment				
Change Assessment Date (If Applicable)	Date when the change impact assessment was conducted				
Assessment Outcome (if Applicable)	Outcome of the assessment (e.g., Change approved, Change Rejected, Further Review Needed)				
Approval Status	The status of approval for the change (Pending, Approved, Rejected)				
Change Status	The current status of the change request or project (In Progress, Completed, Deferred)				
Comments	Any additional notes or comments regarding the request or project				

Change ID	Change Title	Change Description	Request Date	Requested By	Change Owner

6.3.F

Enhance Monitoring with Advanced Reporting: Improve master and reference data monitoring by generating detailed process control reports. These reports should be designed primarily for use by data engineers and quality managers and may include custom reports that monitor specific quality checks, such as completeness and accuracy, at each stage of data processing. For example, a report could track whether critical data fields, like "Customer Address," meet completeness standards during data updates.

A

6.3.G

A

Implement Automated Alerts and Notifications in MDM/RDM Systems: Set up automated alerts within the Master Data Management (MDM) or Reference Data Management (RDM) systems to notify data stewards and data owners of any quality issues with master and reference data. Notifications should be triggered when quality thresholds are breached, for example, If duplicate records are detected in the "Customer ID" field, an alert is sent to the relevant data steward to initiate corrective actions.

RACI Matrix – MRDM Monitoring & Reporting

Indicative

Sr No.	Activity	DGO	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Business Data Analyst	Data Visualization Officer	Data Steward	Data Quality Specialist	IT Team
6.3.A	Establish Key Performance Indicators (KPIs) for Master & Reference Data Management		I	A	R			C	R	
6.3.B	Generate Master & Reference Data Quality Scorecards			A	C	R	R	C	C	
6.3.C	Develop Data Quality Dashboards for Master & Reference Data			A	C	R	R	C	C	
6.3.D	Share Scorecards and Dashboards with the DGO	I		A	R					
6.3.E	Create and Maintain a Change Request Register			A	R			C	C	
6.3.F	Enhance Monitoring with Advanced Reporting		I	A	C			C	R	R
6.3.G	Implement Automated Alerts and Notifications in MDM/RDM Systems		I	A	C			C	R	R

RACI Matrix Legend: R: Responsible A: Accountable C: Consulted I: Informed

Legend: Capability Level

F

Foundational

A

Advanced

VA

Very Advanced

Based on this Activity, the Entity can



Establish KPIs for MRDM

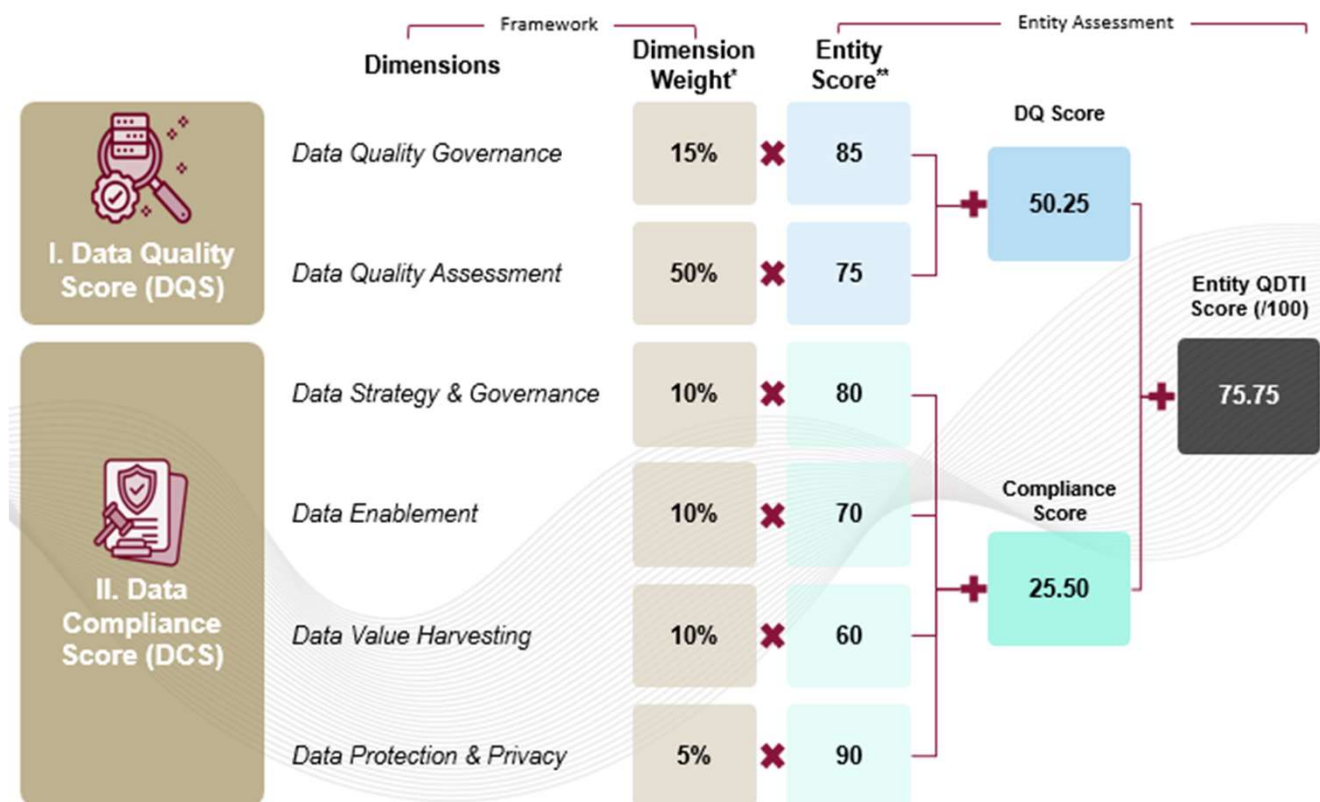


Develop a Data Quality Dashboard and Scorecard

6.4 QDTI

The Qatar Data Trust Index (QDTI) is designed to measure and enhance data integrity across national entities by promoting adherence to governance standards and ensuring the accuracy, reliability, and responsible management of data. This initiative aims to build trust and support a robust digital economy. The overall QDTI score is derived from two dimensions: the Data Quality Score (DQS) and the Data Compliance Score (DCS), which together combine to determine the Entity's QDTI score.

This section outlines the steps needed to gather the necessary information for calculating the DCS.



***Note:** The sum of the weights across all dimensions should add up to 100%

****Note:** This score is calculated by computing the weighted average score of all the sub-dimensions under each respective Dimension (scores are on a scale of 0 to 100)

Figure 20: Illustrative QDTI Scoring Methodology for an Entity

6.4.A

Initiate Master and Reference Data Management Compliance Assessment: The Data Management and Governance Section Head assigns the Data Compliance Specialist to oversee the assessment process. The specialist is responsible for managing the evaluation of the Entity's adherence to the National Master and Reference Data Management Standards.



Key Message / Information

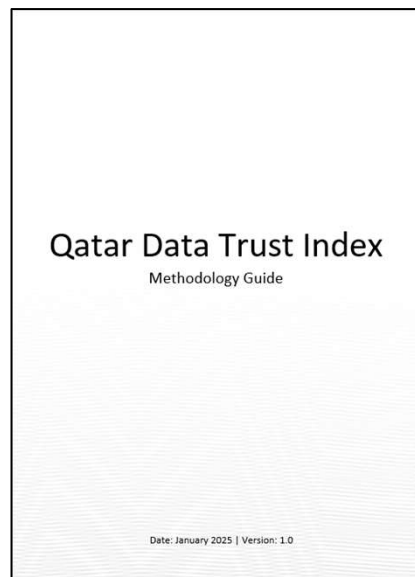
The Master and Reference Data Management Domain falls under the Data Enablement Dimension in the National Data Management Framework and, consequently, within the QDTI Dimension.

6.4.B

Conduct Master and Reference Data Management Assessment: Initiate the self-assessment to evaluate adherence to the National Master and Reference Data Management Domain Standards. This assessment will ensure compliance and accountability while identifying areas for improvement.

 Asset Guidance

The **Qatar Data Trust Index Methodology Guide** can be leveraged to guide the Entity through the steps required to conduct the self-assessment.

Illustrative **Asset Guidance**

The **QDTI Assessment Template** can be used to document compliance with the National Data Standards. This template is intended for use prior to the DGO launching the self-assessment on its platform to calculate the overall QDTI Score.

Illustrative

Domain	Control	Control Description
DQM	DQM.1.1.1 Data Quality Policy	As part of the Data Quality Policy control, the Entity shall create and maintain an enterprise-wide Data Quality Management Policy...
DQM	DQM.1.1.2 Data Quality Policy	As part of the Data Quality Policy control, the Entity shall publish and distribute a Data Quality Management Policy...
DQM	DQM.1.1.3 Data Quality Planning	As part of the Data Quality Planning control, the Entity shall perform initial data quality profiling for the identified Critical Data Elements (CDEs)...
DQM	DQM.1.1.4 Data Quality Planning	As part of the Data Quality Planning control, the Entity shall develop a data quality plan aligned with its data strategy and profiling results...

Specification	Control Specification	Capability Level	Compliance
DQM.1.1.1.1 Data Quality Policy Creation	The Entity shall create a data quality policy in alignment with the NDOAF published by the NPC...	Foundational	Compliant
DQM.1.1.2.1 Data Quality Policy Publication	The Entity shall publish a data quality policy....	Foundational	Compliant
DQM.1.1.3.1 Initial Data Quality Profiling	The Entity shall identify CDEs and perform initial data quality profiling for the identified CDEs....	Foundational	Compliant
DQM.1.1.4.1 Data Quality Plan Design	Based on the initial data profiling results, the Entity shall develop a data quality plan in alignment with the Entity's data strategy...	Foundational	Compliant

**Key Message / Information**

The DGO will launch the self-assessment on its platform and will provide toolkits and host workshops to guide entities through the self-assessment process during the first year.

6.4.C

Submit Master and Reference Data Management Score to DGO. Submit the Master and Reference Data Management Compliance Score to the DGO for evaluation



Key Message / Information

The DGO will calculate the Entity's overall Qatar Data Trust Index (QDTI) score only after the Entity has completed both the Data Quality Score (DQS) assessment and the required self-assessments for all other data management domains to determine the Data Compliance Score (DCS). Once all assessments are completed, the DGO will aggregate the compliance scores across domains to finalize the DCS, which will then be combined with the DQS to determine the Entity's overall QDTI score.

RACI Matrix - QDTI

Indicative

Sr No.	Activity	DGO	Data Management & Security Director	Data Management and Governance Section Head	Data Compliance Specialist	Data Governance Specialist
6.4.A	Initiate Master and Reference Data Management Compliance Assessment		I	A	R	
6.4.B	Conduct Master and Reference Data Management Assessment		I	A	R	C
6.4.C	Submit Master and Reference Data Management Score to DGO	I	I	A	R	I

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Calculate MRDM Compliance Score

6.5 MRDM SLAs

This activity will guide the Entity through the steps to establish Master & Reference Data Management SLAs that specify performance criteria for managing master and reference data. At the foundational level, the Entity will align SLAs with established data quality standards, defining essential criteria such as acceptable quality thresholds, timelines for issue resolution, and escalation processes. For enhanced SLA capabilities, this activity includes implementing automated monitoring and notifications, ensuring that stakeholders are alerted to any SLA breaches in real-time.

Process – MRDM SLAs

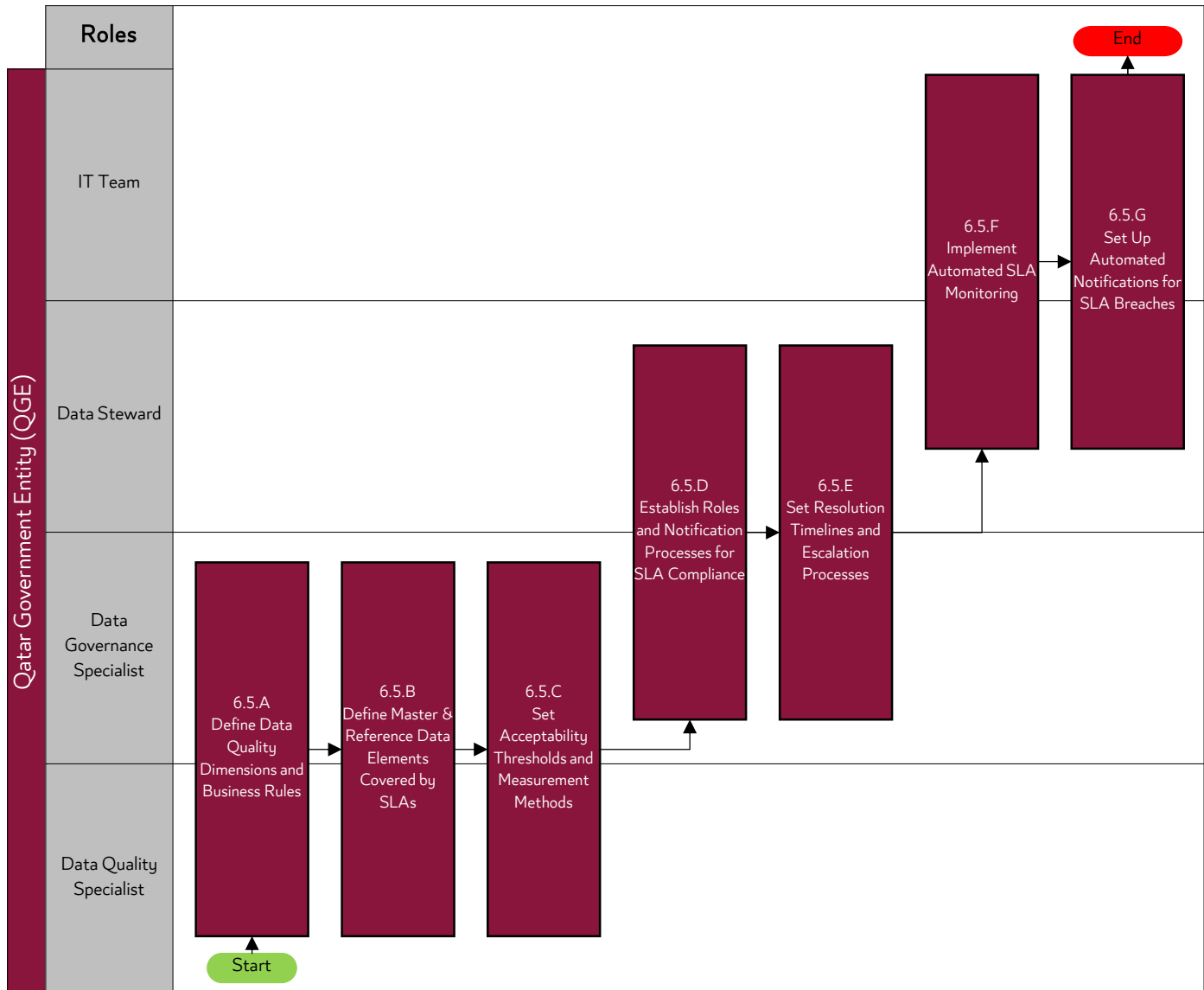


Figure 21: MRDM SLAs Process Map

Legend: Process Step Decision Process Start Process End Document Off-Page Reference

6.5.A

Define Data Quality Dimensions and Business Rules: Specify the data quality dimensions (e.g., accuracy, completeness, timeliness) and business rules for each data attribute covered by the SLAs. Ensure alignment with existing data quality standards to support consistent quality measurement and management.

6.5.B

Define Master & Reference Data Elements Covered by SLAs: Identify the critical master and reference data attributes that require SLA coverage. This includes determining which data attributes are most crucial to business operations and therefore require clearly defined performance criteria.

6.5.C

Set Acceptability Thresholds and Measurement Methods: Define the acceptability thresholds for each data quality dimension, specifying what constitutes an acceptable level of quality for each data attribute. Outline the methods for measuring these thresholds to ensure objective and consistent evaluation.

6.5.D

Establish Roles and Notification Processes for SLA Compliance: Assign data stewards or relevant stakeholders to monitor SLA compliance. Define the notification process for alerting stewards if an SLA threshold is not met, ensuring that appropriate personnel are informed and can act when quality issues arise.

6.5.E

Set Resolution Timelines and Escalation Processes: Define timelines for the resolution of quality issues, along with escalation procedures if SLAs are breached. This step includes specifying deadlines for resolving issues and the process for escalating unresolved issues to higher levels of governance.

6.5.F

Implement Automated SLA Monitoring: Set up automated systems to monitor compliance with SLAs, tracking quality metrics and compliance in real time. This includes automating the measurement of data quality against SLA thresholds, enabling continuous oversight without manual checks.

A

6.5.G

Set Up Automated Notifications for SLA Breaches: Implement automated alerts that notify data stewards and relevant stakeholders when an SLA threshold is breached. This includes setting up automated notifications via email or dashboard alerts to inform stakeholders immediately when data quality issues occur.

RACI Matrix – MRDM SLAs

Indicative

Sr No.	Activity	Data Management and Security Director	Data Management and Governance Section Head	Data Governance Specialist	Data Steward	Data Quality Specialist	IT Team
6.5.A	Define Data Quality Dimensions and Business Rules	I	A	R	C	R	
6.5.B	Define Master & Reference Data Elements Covered by SLAs	I	A	R	C	R	
6.5.C	Set Acceptability Thresholds and Measurement Methods		A	R	C	R	
6.5.D	Establish Roles and Notification Processes for SLA Compliance		A	R	R	C	
6.5.E	Set Resolution Timelines and Escalation Processes		A	R	R	C	
6.5.F	Implement Automated SLA Monitoring	I	A	C	R	C	R
6.5.G	Set Up Automated Notifications for SLA Breaches	I	A	C	R	C	R

RACI Matrix Legend: **R:** Responsible **A:** Accountable **C:** Consulted **I:** Informed

Based on this Activity, the Entity can



Define MRDM SLAs

7. Appendix



8.1. Templates and Frameworks

Indicative

Asset No.	Asset Name	Description	File Name
1	Master and Reference Data Management Policy Template	Template containing a sample table of contents the Entity can refer to when drafting its policy.	NPC_DGO_QDKC_MRDM_Policy Template
2	Data Management Policy Communication Email Template	Template containing a sample email the Entity can use to announce the publishing of the new policy.	NPC_DGO_QDKC_General_Data Management Policy Communication Email Template
3	CDE Prioritization Framework	Framework that provides a weighted criteria to help the Entity prioritize CDEs..	NPC_DGO_QDKC_MRDM_CDE Prioritization Framework
4	Initiative Card Template	Template used to outline and track the details of a specific initiative.	NPC_DGO_QDKC_General_Initiative Card Template
5	Data Management Domain Plan Template	Template that helps the Entity capture an overview of identified initiatives, detail their sub-activities, owners, stakeholders, and activity timeframe.	NPC_DGO_QDKC_General_Data Management Domain Plan Template
6	Data Catalog Template	Template containing the dataset inventory template, attributes inventory template, and relational database metadata schema.	NPC_DGO_QDKC_DCMM_Data Catalog Template
7	MRDM Rules Template	Template containing 3 sheets to support the Entity in defining attribute trust, similarity, and survivorship rules.	NPC_DGO_QDKC_MRDM_MRDM Rules Template
8	MRD Requirements Templates	Template that helps the Entity document overall requirements and detail specific MRD Requirements.	NPC_DGO_QDKC_MRDM_MRDM Requirements Template
9	MDM Tool Selection Criteria Template	Template containing key factors for evaluating and selecting MDM tools, covering functionality, integration, cost and other factors.	NPC_DGO_QDKC_MRDM_Tool Selection Criteria Template

Asset No.	Asset Name	Description	File Name
10	Data Management KPIs Register Template	Template containing a table that helps the Entity define KPIs.	NPC_DGO_QDKC_General_Data Management KPIs Register Template
11	Data Quality Scorecard	Template containing performance metrics for evaluating data quality against standards, helping track improvements and communicate status to stakeholders.	NPC_DGO_QDKC_DQ_Data Quality Scorecard
12	Data Management Change Management Template	Template containing a training plan template that the Entity can use to design a training plan for promoting and upskilling its staff.	NPC_DGO_QDKC_General_Data Management Change Management Template
13	Qatar Data Trust Index Methodology Guide	The Qatar Data Trust Index Methodology Guide can be leveraged to guide the Entity through the compliance self-assessment.	22012025_NPC_QDTI Methodology Guide_V1.0
14	QDTI Assessment Template	The QDTI Assessment Template can be used to document compliance with the National Data Standards.	23012025_NPC_QDTI Assessment Template_V1.0



المجلس الوطني للتخطيط
National Planning Council

دولة قطر • State of Qatar