

# National Data Management Guidelines

## Data Quality Management



## Table of Contents

### 1. Data Quality 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 Data Quality Management Process Map

### 2. Data Quality Policy

- Overview
- 2.1. Data Quality Policy Creation
- 2.2. Data Quality Policy Publication

### 3. Data Quality Planning

- Overview
- 3.1. Data Profiling
- 3.2. Data Quality Plan

### 4. Data Quality Implementation

- Overview
- 4.1. Data Quality Rules & Thresholds
- 4.2. Data Quality Tool Implementation

### 5. Data Quality Issue Management

- Overview
- 5.1. Data Quality Issues Management and Remediation Process

### 6. Data Quality Performance Management

- Overview
- 6.1. Data Quality KPIs
- 6.2. Data Quality Monitoring & Reporting
- 6.3. QDTI
- 6.3. Data Quality SLAs

### 7. Appendix

- 7.1. Templates and Frameworks

# 1. Data Quality Management Introduction



## 1.1 Domain Overview

Data quality refers to the overall condition of data as assessed by criteria such as accuracy, completeness, consistency, validity, and timeliness. Evaluating data quality levels enables Entities to identify errors and inconsistencies that require resolution, ensuring that the data housed in IT systems meets the expectations and needs of its users.

Data Quality Management (DQM) includes the planning, implementation, and oversight of activities that employ quality management techniques to ensure data is fit for use and aligns with the requirements of data consumers. Effective DQM not only enhances the integrity of data but also promotes trust among stakeholders who rely on this information for decision-making.

### Key Components of the Data Quality Management Domain:

## 2. Data Quality Policy

The Data Quality Policy is the cornerstone of an Entity's data governance framework, ensuring that data assets are managed effectively and consistently. This policy outlines the frameworks, principles, and practices necessary to maintain high-quality data, supporting the Entity's strategic objectives.

- **Data Quality Policy Creation:** Creating a Data Quality Policy involves defining the principles, guidelines, and standards that ensure data accuracy, completeness, consistency, reliability, and timeliness. This process requires collaboration across various stakeholders to align the policy with the Entity's goals and regulatory requirements.
- **Data Quality Policy Publication:** Once the Data Quality Policy is created, it must be formally documented and communicated to all relevant stakeholders. This publication process includes disseminating the policy through official channels, ensuring accessibility, and providing training and resources to support adherence.

## 3. Data Quality Planning

Data Quality Planning is a proactive approach to managing data quality, focusing on identifying potential data issues and establishing measures to address them. It involves detailed analysis and strategic planning to integrate data quality management into the Entity's operations.

- **Data Profiling:** Data Profiling is the initial step in data quality planning, involving the examination of data sources to understand their structure, content, and interrelationships. This process helps identify data anomalies, inconsistencies, and potential quality issues that need to be addressed.
- **Data Quality Plan:** The Data Quality Plan outlines specific actions and resources required to address identified data quality issues. This plan includes setting objectives, defining key performance indicators (KPIs), allocating responsibilities, and establishing timelines to ensure continuous improvement in data quality.

#### 4. Data Quality Implementation

Data Quality Implementation translates the relevant data management policies and plans into actionable practices. It involves deploying tools, defining rules, and establishing processes to ensure that data quality standards are met throughout the data lifecycle.

- **Data Quality Rules & Thresholds:** Defining Data Quality Rules and Thresholds involves setting specific criteria for data quality, such as acceptable levels of accuracy, completeness, and consistency. These rules and thresholds guide data validation processes and ensure that data meets the required standards.
- **Data Quality Tool Implementation:** Implementing Data Quality Tools involves selecting and deploying software solutions that automate data quality tasks, such as data cleansing, validation, and monitoring. These tools support the enforcement of data quality rules and facilitate ongoing quality management.

#### 5. Data Quality Issue Management

Data Quality Issue Management focuses on identifying, tracking, and resolving data quality issues. This process ensures that data anomalies are promptly addressed to maintain data integrity and reliability.

- **Data Quality Issues Management and Remediation Process:** The Data Quality Issues Management and Remediation Process outlines the steps for logging and prioritizing data quality issues, analyzing their root causes, and implementing corrective and preventive actions. This process ensures systematic handling of data quality problems to prevent recurrence and improve overall data quality.

#### 6. Data Quality Performance Management

Data Quality Performance Management involves monitoring, measuring, and reporting on data quality to ensure continuous improvement. This process integrates performance metrics and feedback mechanisms to sustain high data quality standards.

- **Data Quality KPIs:** Data Quality Key Performance Indicators (KPIs) are metrics used to assess the effectiveness of data quality management efforts. KPIs provide quantifiable measures of data quality, such as error rates, data completeness, and timeliness, guiding decision-making and improvement initiatives.
- **Data Quality Monitoring & Reporting:** Data Quality Monitoring and Reporting involves ongoing surveillance of data quality metrics and the generation of reports to communicate performance to stakeholders. This process helps identify trends, track progress, and highlight areas needing attention.
- **QDTI:** Evaluating the Entity's adherence to the National Data Quality Management Standards and the Data Quality Assessment results to determine the Data Quality Score (DQS), a key component of the Qatar Data Trust Index (QDTI) score.
- **Data Quality SLAs:** Data Quality Service Level Agreements (SLAs) define the expected quality standards and performance levels for data management. SLAs establish clear expectations between data providers and consumers, ensuring accountability and alignment with organizational goals.

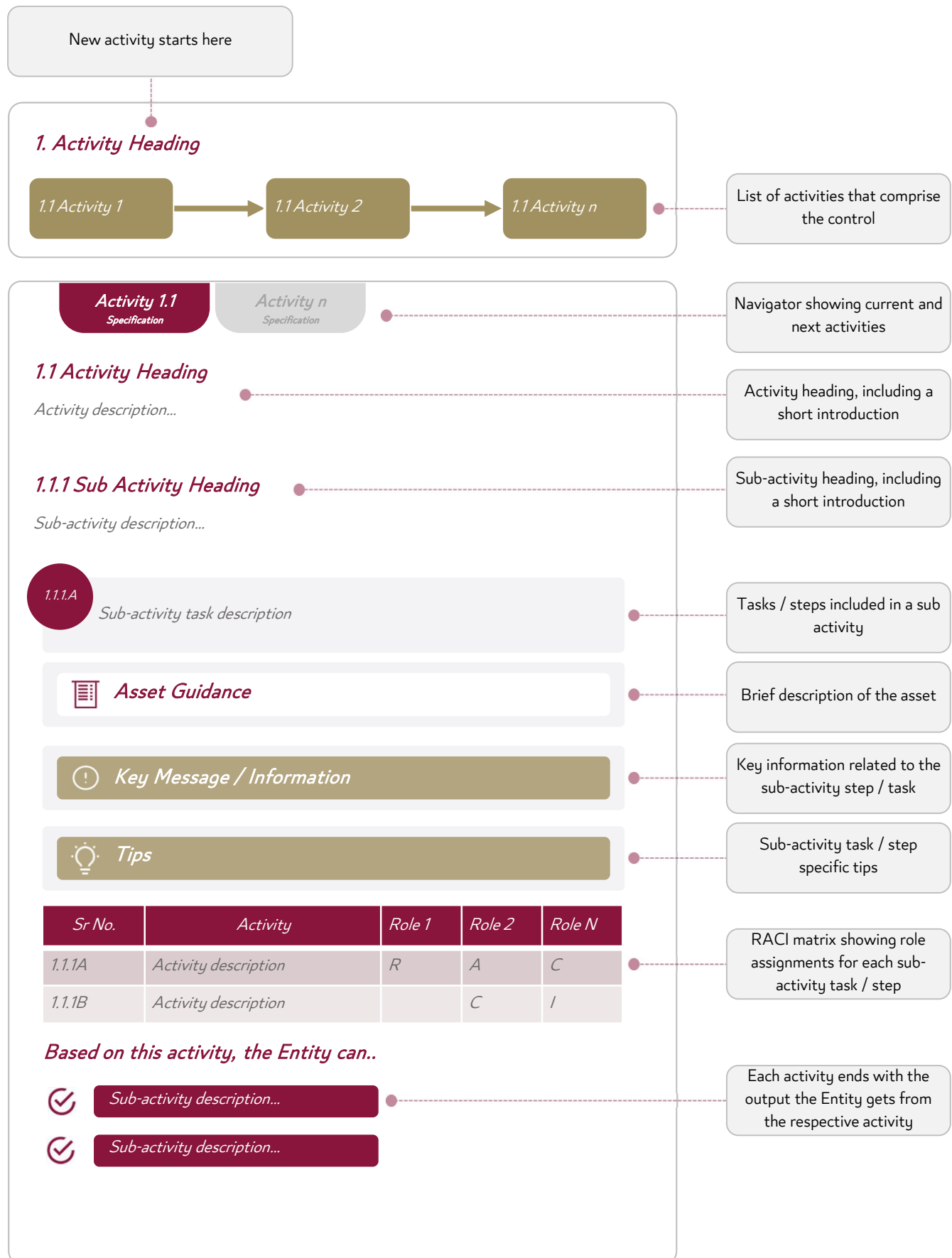
## 1.2 Target Audience

This Data Quality Domain handbook is intended for business users and the Entity's Data Quality working group including the Data Quality Specialist, Data Owners, Stewards, and Analysts who are responsible for defining the data quality requirements across the different data domains and developing appropriate measures for assessing and improving the data quality levels within the Entity. This is demonstrated through detailed step by step processes, activities, and examples, that can be followed to ensure that the Entity is able to leverage accurate and reliable data to drive its business and create value.

### 1.3 Abbreviations and Glossary

| Abbreviation/Term       | Definition / Long Name                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>AI</b>               | Artificial Intelligence (AI) is the simulation of human intelligence by machines, particularly computer systems. It enables computers to perform tasks like learning, reasoning, problem-solving, perception, and language understanding. AI includes subfields such as machine learning, natural language processing, robotics, and computer vision, aiming to create systems that can learn and adapt autonomously.                                                  |
| <b>CDE</b>              | Critical Data Elements (CDEs) are essential pieces of information crucial to an Entity's operations, decision-making, and compliance. They significantly impact business processes and include data such as customer details, product information, financial records, and employee data. Maintaining the quality of CDEs is essential for operational efficiency, supporting informed decisions, and achieving regulatory compliance.                                  |
| <b>DQ</b>               | Data quality (DQ) refers to the condition of data based on factors such as accuracy, consistency, completeness, reliability, and relevance. High-quality data is essential for effective decision-making, operational efficiency, and compliance with regulatory requirements. It indicates that the information used by an Entity is trustworthy and fit for its intended purposes.                                                                                   |
| <b>Data Stewardship</b> | Data Stewardship refers to the assignment of dedicated individuals or teams (data stewards) responsible for overseeing and maintaining the quality of data within an Entity. Data stewards ensure compliance with established data quality standards by monitoring, validating, and improving data accuracy, consistency, completeness, and validity. Their role is critical in upholding governance practices and ensuring data remains reliable and fit for purpose. |
| <b>DQM</b>              | Data Quality Management (DQM) ensures an Entity's data is accurate, consistent, complete, and reliable. It involves practices like data assessment, cleansing, profiling, and governance to maintain high data standards. Effective DQM enhances decision-making, regulatory compliance, and operational efficiency.                                                                                                                                                   |
| <b>KPI</b>              | A Key Performance Indicator (KPI) is a measurable value that indicates how effectively an Entity is achieving its key business objectives. KPIs are used to evaluate the success of an Entity, department, or project in reaching targets. They provide actionable insights and help in monitoring progress, identifying areas for improvement, and making informed decisions.                                                                                         |
| <b>Metadata</b>         | Information about the data collected, this includes information about technical and business processes, data rules and constraints and logical and physical data structures                                                                                                                                                                                                                                                                                            |
| <b>NDQF</b>             | National Data Quality Framework                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <b>NPC</b>              | National Planning Council                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>NQAF</b>             | National Quality Assurance Framework                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>QDKC</b>             | Qatar Data Knowledge Case                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>RACI</b>             | Responsible, Accountable, Consulted, Informed                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>SME</b>              | Subject Matter Expert                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

## 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 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.                                                                                                                                                                                                                                                                                                                                                                                                          |
| Business Data Analyst                        | The Business Data Analyst is responsible for analyzing data from multiple sources to provide actionable insights that drive strategic and operational decisions within the Entity. This role involves designing reports and dashboards to effectively communicate findings, collaborating with cross-functional teams to ensure data accuracy and relevance, and aligning analytics with business objectives. The Analyst plays a key role in enabling data-driven decision-making and identifying opportunities to improve processes and outcomes.                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| 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.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

| Stakeholder                | Role Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 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. |
| Data Compliance Specialist | The Data Compliance Specialist ensures the Entity's data practices comply with national regulations, organizational policies, and international standards. This role manages data privacy, security, and governance compliance, conducts audits, mitigates non-compliance risks, and promotes awareness of regulatory requirements. The specialist works with stakeholders to embed compliance controls into data processes, ensuring adherence to legal and ethical standards.                   |
| Data Engineer              | The Data Engineer transforms data into a format that can be easily analyzed and utilized. The Data Engineer works closely with other stakeholders to ensure data quality is maintained and implemented systems adhere to the highest possible standards.                                                                                                                                                                                                                                          |
| 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.                                                                                                         |
| 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.                                                                               |
| 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.                                                                                                                |

| Stakeholder   | Role Description                                                                                                                                                                                                                                                                                                                                    |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data Stewards | <p>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.</p> |

## 1.6 Key Success Factors

**Prioritize Critical Data:** A successful data management program should prioritize data elements that are most relevant to the organization's core business processes, reports, and analytics. This focus ensures that efforts are concentrated on data that directly impacts the entity's success, enabling more efficient resource allocation and better support for strategic objectives.

**Define Data Quality Standards:** Establishing and tracking target scores and thresholds for various data quality dimensions is essential. These standards should be reported to leadership as key performance indicators (KPIs) to provide visibility into the health of data assets and to drive accountability for data quality across the organization.

**Integrate Quality Throughout the Data Lifecycle:** Data quality should be maintained at every stage of the data lifecycle, from acquisition to retirement. Implementing data quality gates at critical points ensures that high standards are upheld and that any issues are identified and addressed promptly.

**Address Root Causes:** To achieve lasting improvements in data quality, it is crucial to fix issues at their source in core systems. This approach prevents recurring problems and reduces the need for continuous downstream corrections, ensuring a more stable and reliable data environment.

**Monitor Continuously:** Continuous measurement and reporting of data quality metrics for each defined dimension are necessary to maintain sustainable quality. Regular monitoring helps in identifying trends, detecting anomalies, and ensuring that data remains fit for purpose over time.

**Promote Awareness:** Raising awareness about the importance of good data quality is vital for fostering a culture of data excellence. Transparent communication about data quality improvements and their positive outcomes helps to build trust and encourages ongoing commitment to data quality initiatives across the organization.

The following chapters provide detailed guidance in each of these areas, offering step-by-step instructions to establish a strong data quality management program and advance data maturity within the Entity.

## 1.7 Compliance with National Data Standards

| National Data Standards                     |                                                                               | QDKC                                                     |
|---------------------------------------------|-------------------------------------------------------------------------------|----------------------------------------------------------|
| Control                                     | Specification                                                                 | Activity                                                 |
| DQM.1.1 Data Quality Policy                 | DQM.1.1.1 Data Quality Policy Creation <span>F</span>                         | 2.1 Data Quality Policy Creation                         |
|                                             | DQM.1.1.2 Data Quality Policy Publication <span>F</span>                      | 2.2 Data Quality Policy Publication                      |
| DQM.1.2 Data Quality Planning               | DQM.1.2.1 Data Profiling <span>F</span>                                       | 3.1 Data Profiling                                       |
|                                             | DQM.1.2.2 Data Quality Plan <span>F</span>                                    | 3.2 Data Quality Plan                                    |
| DQM.2.1 Data Quality Implementation         | DQM.2.1.1 Data Quality Rules & Thresholds <span>F</span>                      | 4.1 Data Quality Rules & Thresholds                      |
|                                             | DQM.2.1.2 Data Quality Tool Implementation <span>F</span>                     | 4.2 Data Quality Tool Implementation                     |
|                                             | DQM.2.1.3 Data Quality Tool Implementation <span>A</span>                     |                                                          |
|                                             | DQM.2.1.4 Data Quality Tool Implementation <span>VA</span>                    |                                                          |
| DQM.3.1 Data Quality Issue Management       | DQM.3.1.1 Data Quality Issues Management & Remediation Process <span>F</span> | 5.1 Data Quality Issues Management & Remediation Process |
| DQM.3.2 Data Quality Performance Management | DQM.3.2.1 Data Quality KPIs <span>F</span>                                    | 6.1 Data Quality KPIs                                    |
|                                             | DQM.3.2.2 Data Quality Monitoring & Reporting <span>F</span>                  | 6.2 Data Quality Monitoring & Reporting                  |
|                                             | DQM.3.2.3 Data Quality Monitoring & Reporting <span>A</span>                  |                                                          |
|                                             | DQM.3.2.4 Data Quality SLAs <span>F</span>                                    | 6.3 Data Quality SLAs                                    |
|                                             | DQM.3.2.5 Data Quality SLAs <span>A</span>                                    |                                                          |

Legend: Capability Level



Foundational

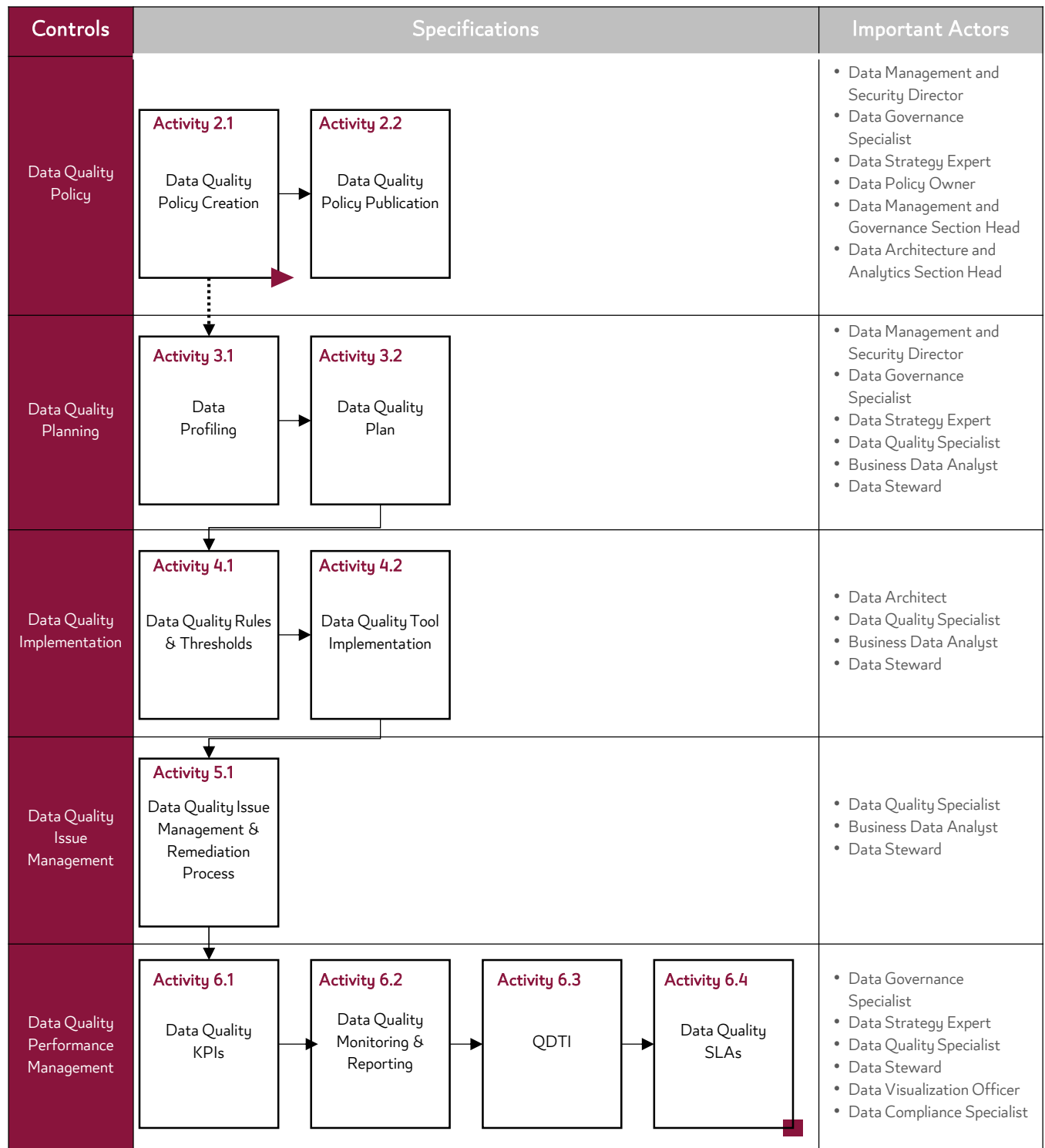


Advanced



Very Advanced

## 1.8 Data Quality Management Process Map



## 2. Data Quality Policy



## 2. Data Quality Policy

### What is a Data Quality Policy and why does the Entity need it?

A data quality policy is a set of guidelines designed to ensure that an Entity's data is accurate, consistent, complete, and reliable. This policy outlines the frameworks, principles and practices for managing data quality across the Entity, including the identification of critical data elements, the development of data quality rules, data profiling, quality assessments, issue management, and performance monitoring.

Data Quality Policy Key Benefits:

- **Promotes Data Integrity:** High-quality data is crucial for accurate decision-making, reporting, and analytics. A data quality policy helps maintain the integrity of data by setting standards for its collection, storage, and usage.
- **Improves Compliance:** Many industries are subject to regulations that require strict data management practices. A data quality policy helps ensure compliance with these regulatory requirements, avoiding legal issues and penalties.
- **Enhances Efficiency:** By establishing clear procedures for data management, the policy helps streamline data handling processes, reducing errors and redundancies. This leads to more efficient operations and better resource utilization.
- **Supports Decision-Making:** Reliable data is the foundation of sound business decisions. A data quality policy ensures that decision-makers have access to accurate and timely data, enhancing the quality of strategic and operational decisions.
- **Facilitates Trust:** Consistently high data quality builds trust among stakeholders, including employees, customers, and partners. It demonstrates the Entity's commitment to data governance and responsible data management.
- **Enables Effective Data Integration:** As Entities integrate data from various sources, a data quality policy ensures that the data remains consistent and accurate, enabling effective data integration and utilization.

The remaining sections of this chapter will detail the below processes to guide the execution of the Data Quality Policy:



## Guiding Principles for Data Quality Policy

1. **Appoint a Single Owner:** Designate a single owner or data steward for data policies to ensure accountability and clear leadership in data governance.
2. **Maintain a Data Policy Repository:** Publish and maintain a centralized repository of all data management related policies that is easily accessible to all employees, promoting transparency and awareness.
3. **Annual Review and Updates:** Review data policies at least annually to ensure they remain valid and relevant, using disciplined version control to track changes and updates.
4. **Stakeholder Engagement:** Involve key stakeholders from different departments in the policy development and review process to ensure diverse perspectives and needs are considered.
5. **Training and Awareness:** Implement regular training sessions to educate employees about data policies and their importance, fostering a culture of data quality.
6. **Compliance and Standards:** Ensure that data policies align with relevant regulatory requirements and industry standards, facilitating compliance and risk management.
7. **Data Quality Metrics:** Establish clear metrics for measuring data quality and the effectiveness of data policies, enabling ongoing monitoring and assessment.
8. **Feedback Mechanism:** Create a feedback loop where employees can report issues or suggest improvements to data policies, encouraging continuous improvement.
9. **Documentation of Procedures:** Clearly document procedures for data collection, processing, and reporting to ensure consistency and support adherence to policies.
10. **Change Management:** Develop a change management process for implementing updates to data policies, ensuring smooth transitions and minimal disruption.
11. **Integration with Business Processes:** Ensure data quality policies are integrated into business processes and workflows, making them a fundamental part of organizational operations.
12. **Technology Support:** Leverage technology tools to support data quality initiatives, such as data profiling, cleansing, and monitoring solutions.

## 2.1 Data Quality Policy Creation

The Data Quality Policy should be designed to uphold the highest standards of data quality across all systems, platforms, and data sources within the Entity. It must be aligned with the National Quality Assurance Framework (NQAF) as set forth by the National Planning Council (NPC) to ensure adherence to national standards and best practices. This policy is essential not only for certifying that data meets the rigorous requirements for statistical and analytical purposes but also for driving continuous improvements in data quality across the Entity. By establishing a clear and comprehensive framework, the Data Quality Policy will enable the Entity to maintain reliable, accurate data that supports informed decision-making, improves operational efficiency, and promotes a culture of data excellence throughout all departments.

### Process - Data Quality Policy Creation

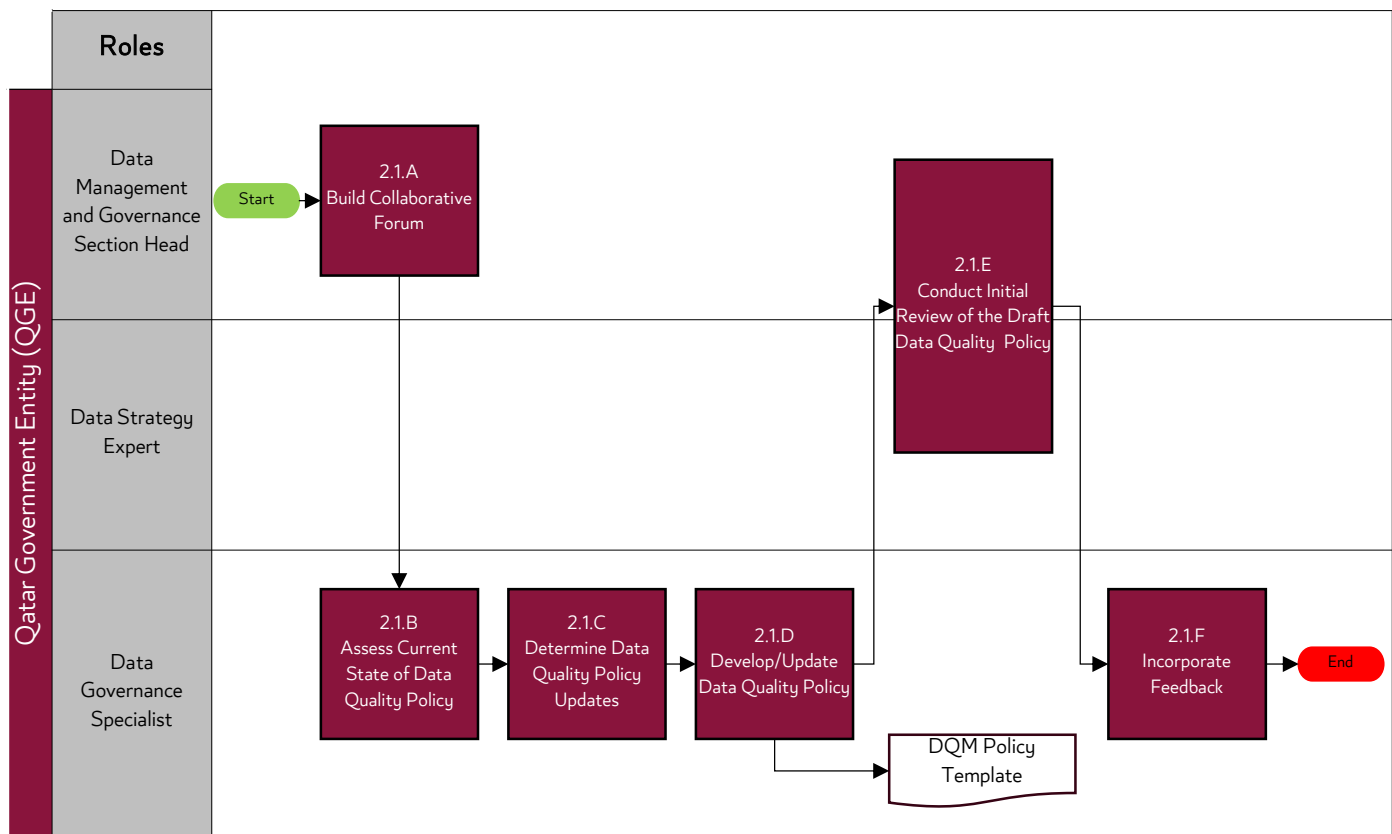


Figure 1: Data Quality Policy Creation Process Map

**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 State of Data Quality Policy:*** Assess the current state of the Entity's data quality 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 Data Quality 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 Data Quality Policy:** Using the Data Quality Policy template, the Data Governance Specialist develops or updates the document as a preliminary policy.



## 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.

### Asset Guidance

The **Data Quality Policy Template** provides a sample table of contents for drafting the 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                 | Data Quality Policies.....#                          |
| 2.1               | Data Quality Dimensions.....#                        |
| 2.2               | Critical Data Elements Identification.....#          |
| 2.3               | Data Quality Rules and Thresholds.....#              |
| 2.4               | Data Profiling and Quality Assessment.....#          |
| 2.5               | Data Quality Escalations and Issue Management.....#  |
| 2.6               | Data Quality Remediation Process.....#               |
| 2.7               | Performance Management and Monitoring.....#          |
| 2.8               | Key Policy Performance Indicators.....#              |
| 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 Data Quality Policies |                                                                                                                                                                                                                                       |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2.1                     | Data Quality Dimensions                                                                                                                                                                                                               |
| 1.                      | All organizational data shall be assessed and managed across defined data quality dimensions, including accuracy, completeness, consistency, timeliness, validity, and uniqueness, to ensure it meets business and operational needs. |
| 2.                      | Data accuracy shall be prioritized to ensure that all data values are represented correctly, reducing the risk of errors in decision-making and reporting.                                                                            |
| 3.                      | Completeness of data shall be monitored to guarantee that all required fields and records are captured and maintained for critical business processes.                                                                                |
| 4.                      | Consistency checks shall be enforced across datasets to eliminate discrepancies between related data values, ensuring alignment within and across systems.                                                                            |
| 5.                      | Data timeliness shall be managed to ensure data is available when needed for decision-making and operations, with clear timelines for updates and validation.                                                                         |
| 6.                      | Validity of data shall be ensured by aligning values with established business rules, data models, and reference datasets to maintain logical correctness.                                                                            |
| 7.                      | Data uniqueness shall be maintained to avoid duplication of records, ensuring that each data item is represented only once within the relevant datasets.                                                                              |
| 8.                      | Data quality metrics and thresholds for each dimension shall be established, monitored, and regularly reviewed to ensure compliance with organizational standards and regulatory requirements.                                        |
| 9.                      | Regular data quality audits shall be conducted to assess adherence to the defined dimensions and address gaps, with corrective actions documented and tracked to resolution.                                                          |

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

2.1.E

**Conduct Initial Review of Draft Data Quality 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 – Data Quality Policy Creation

*Indicative*

| Sr No. | Activity                                            | Data Management and Security Director | Data Management and Governance Section Head | Data Strategy Expert | Data Governance Specialist |
|--------|-----------------------------------------------------|---------------------------------------|---------------------------------------------|----------------------|----------------------------|
| 2.1.A  | Build Collaborative Forum                           | A                                     | R                                           |                      | C                          |
| 2.1.B  | Assess Current State of Data Quality Policy         |                                       | A                                           | C                    | R                          |
| 2.1.C  | Determine Data Quality Policy Updates               |                                       | A                                           | C                    | R                          |
| 2.1.D  | Develop/Update Data Quality Policy                  |                                       | A                                           | C                    | R                          |
| 2.1.E  | Conduct Initial Review of Draft Data Quality Policy |                                       | R/A                                         | R                    | I                          |
| 2.1.F  | Incorporate Feedback                                |                                       | A                                           |                      | R                          |

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

Based on this Activity, the Entity can



Build Collaborative Forum of SMEs



Draft Data Quality Policy

## 2.2 Data Quality Policy Publication

This activity will outline the steps to publish the data quality policy.

### Process - Data Quality Policy Publication

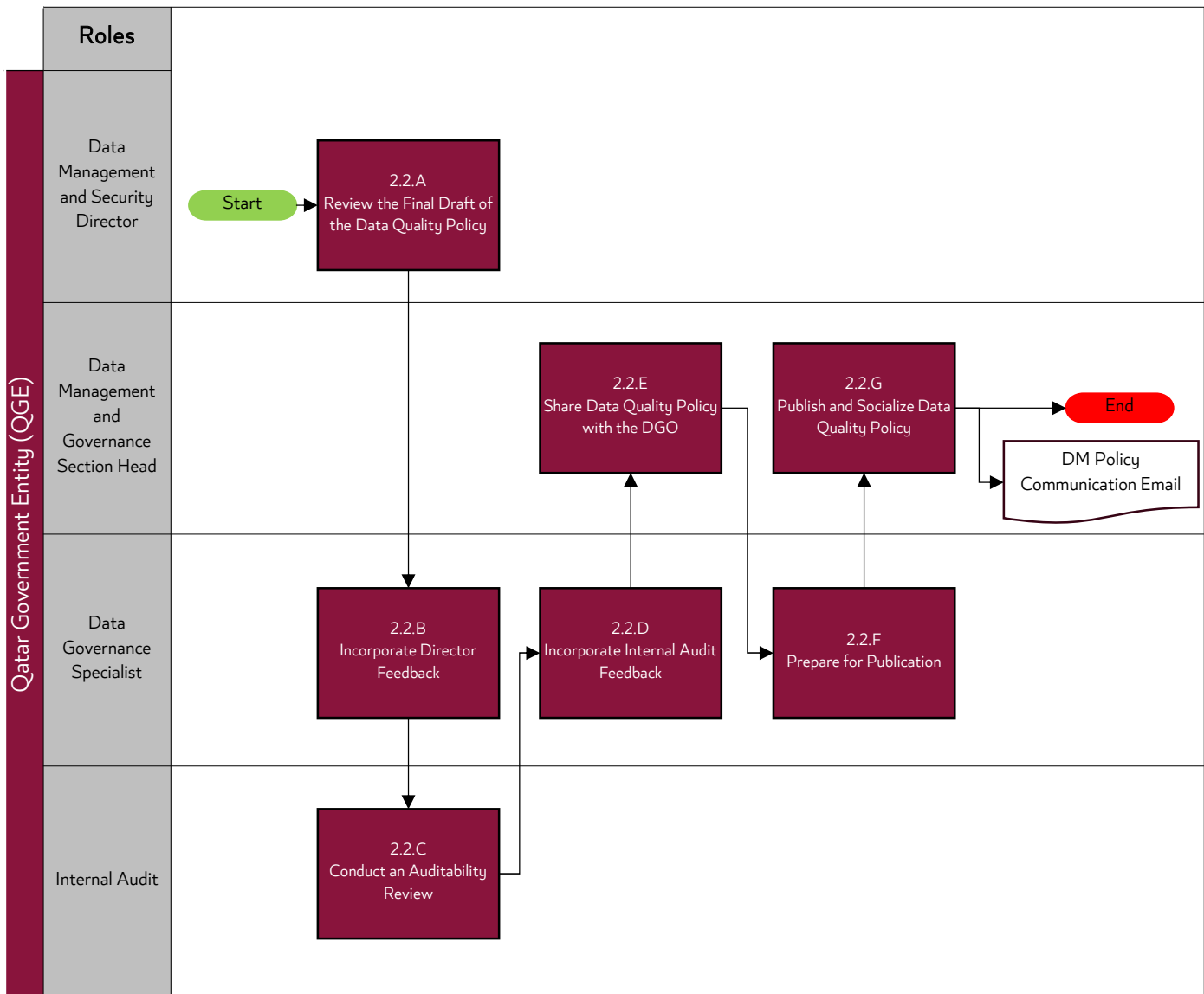


Figure 2: Data Quality Policy Publication Process Map

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

#### 2.2.A

***Review Final Draft of Data Quality Policy for Strategic Alignment:*** The Data Management and Governance Department Director reviews the final draft of the policy 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 Governance Department 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 Data Quality Policy with the DGO:*** The Entity must submit the finalized draft of the Data Quality Policy to the DGO as well as other relevant stakeholders for their review and feedback. This step ensures that the policy aligns with national 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 Data Quality Policy.** Publish the finalized data quality 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 that can be used to notify staff about the publication of the policy and provide key information.

*Illustrative*

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Subject:</b> [IMPORTANT] New [Domain Name] Policy Published – Please Review</p> <p>Dear [Recipient Name/Team],</p> <p>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.</p> <p><b>What is the Policy About?</b></p> <p>The [Domain Name] Policy governs the management, oversight, and compliance of [specific domain] activities across the entity. It outlines critical guidelines and procedures, including:</p> <ul style="list-style-type: none"> <li>• [Key Point 1, e.g., "Data Management and Organizational Structure"]</li> <li>• [Key Point 2, e.g., "Version control and security measures"]</li> <li>• [Key Point 3, e.g., "Metadata quality and governance"]</li> </ul> <p>This policy has been developed in collaboration with the Data Committee and relevant stakeholders, ensuring compliance with the national data management standards.</p> <p><b>How Does This Affect You?</b></p> <ul style="list-style-type: none"> <li>• <b>Action Required:</b> 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.</li> <li>• <b>Changes to Processes:</b> [If applicable, briefly describe any changes to current workflows or responsibilities as outlined by the policy.]</li> </ul> <p><b>Where Can You Access the Policy?</b></p> <p>The policy is available on our internal [intranet/portal] and can be accessed through the following link:<br/>[Link to Policy Document]</p> <p><b>Next Steps</b></p> <ul style="list-style-type: none"> <li>• <b>Review Deadline:</b> Please ensure that you and your team have reviewed the policy by [Date].</li> <li>• <b>Training/Workshops:</b> [If applicable] We will be hosting a [training/workshop] on [Date] to assist teams in understanding and applying the new policy.</li> <li>• <b>Ongoing Support:</b> For questions, feedback, or support, please contact [Support Contact] at [email/phone number].</li> </ul> <p>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.</p> <p>Best regards,<br/>[Your Name]<br/>[Your Position]<br/>[Your Department]<br/>[Contact Information]</p> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

## Process Risks: Develop & Publish Data Quality Management Policies

| 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 data quality 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 data quality policy. |

## RACI Matrix – Data Quality 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 the Data Quality 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 Data Quality Policy with the DGO            | I   | A                                     | R                                           |                            |                      |                |
| 2.2.F  | Prepare for Publication                           |     |                                       | A                                           | R                          | C                    |                |
| 2.2.G  | Publish and Socialize Data Quality Policy         |     | A                                     | R                                           |                            | C                    |                |

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

Based on this Activity, the Entity can



Publish Data Quality Policy

### 3. Data Quality Planning



### 3. Data Quality Planning

#### What is Data Quality Planning and why does the Entity need it?

Data Quality Planning is the process of developing a structured plan to ensure data meets the required quality standards. Based on data profiling results, this plan outlines activities, milestones, and owners responsible for improving data quality. It addresses interdependencies, risks, and assumptions, and includes a timeline and cost for implementation.

Data quality planning aims to systematically improve and maintain high data quality, which is essential for accurate reporting, analytics, and decision-making. A well-defined plan ensures that data quality initiatives are aligned with organizational goals, efficiently managed, and continuously monitored to address gaps and challenges.

#### Data Quality Plan Key Benefits:

- **Promotes Proactive Management:** A robust data quality plan allows Entities to identify potential data quality issues before they arise, enabling proactive measures that prevent costly problems down the line.
- **Establishes Clear Standards:** By defining quality metrics and benchmarks, data quality planning creates a clear framework for what constitutes high-quality data, ensuring consistency in data management practices across the Entity.
- **Enhances Resource Planning:** Effective planning helps Entities allocate resources efficiently, ensuring that time, personnel, and tools are dedicated to maintaining and improving data quality where it is most needed.
- **Improves Stakeholder Communication:** A comprehensive data quality plan fosters better communication among stakeholders by clarifying roles, responsibilities, and expectations regarding data quality, which enhances collaboration across departments.
- **Supports Continuous Improvement:** With a structured approach to data quality, Entities can regularly assess and refine their practices, promoting a culture of continuous improvement and adaptation to changing business needs.
- **Boosts Data-Driven Decision-Making:** By ensuring that data quality is prioritized, Entities enhance their ability to make informed, data-driven decisions, leading to better strategic outcomes and competitive advantages.

The remaining sections of this chapter will detail the below processes to guide the execution of Data Quality Planning:



## Guiding Principles for Data Quality Planning

1. **Define Clear Objectives:** Establish specific goals for data quality that align with the overall business strategy. Identify what high-quality data means for the Entity and how it supports decision-making.
2. **Identify Critical Data Elements:** Focus on critical data elements that drive business processes. Prioritize these elements to ensure that the most impactful data is monitored and maintained rigorously.
3. **Develop Quality Metrics and Benchmarks:** Create clear, measurable metrics to assess data quality. Metrics may include accuracy, completeness, consistency, and validity. Establish benchmarks to gauge performance over time.
4. **Engage Stakeholders:** Involve all relevant stakeholders, including data owners, users, and IT staff, in the planning process. Ensure that roles, responsibilities, and expectations regarding data quality are clearly communicated.
5. **Implement Proactive Monitoring:** Establish processes for continuous monitoring and assessment of data quality. Use automated tools where possible to detect issues early and implement preventive measures.
6. **Allocate Resources Effectively:** Assess resource requirements for maintaining data quality, including personnel, tools, and technology. Allocate resources based on data importance and identified risk areas.
7. **Foster a Culture of Quality:** Promote a culture that values data quality across the Entity. Provide training and support to employees, encouraging them to take ownership of data quality in their roles.
8. **Adapt and Evolve:** Recognize that data quality needs may change over time. Establish a framework for regular review and adaptation of the data quality plan to address new challenges and opportunities.
9. **Integrate with Existing Processes:** Ensure that data quality planning is integrated with other business processes and initiatives, such as compliance, risk management, and performance improvement efforts.
10. **Communicate Progress and Results:** Regularly share updates on data quality initiatives and results with stakeholders. Highlight successes and areas for improvement to maintain engagement and support.

### 3.1 Data Profiling

Data profiling is a critical process for evaluating the quality and relevance of data within the Entity. It involves assessing the accuracy, completeness, and consistency of data to ensure it meets the needs of the business. Through data profiling, the Entity gains deeper insights into the integrity of its data, identifying potential issues and areas for improvement. This process also helps prioritize data elements based on their business impact, enabling informed decision-making and effective resource allocation. By utilizing established data quality dimensions, data profiling provides a structured approach to understanding and enhancing the overall value of the Entity's data assets.

#### 3.1.1 Define Critical Data Elements (CDEs)

This activity will guide the Entity on the steps to identify and prioritize critical data elements and document them in the data catalog.

#### Process - Define Critical Data Elements (CDEs)

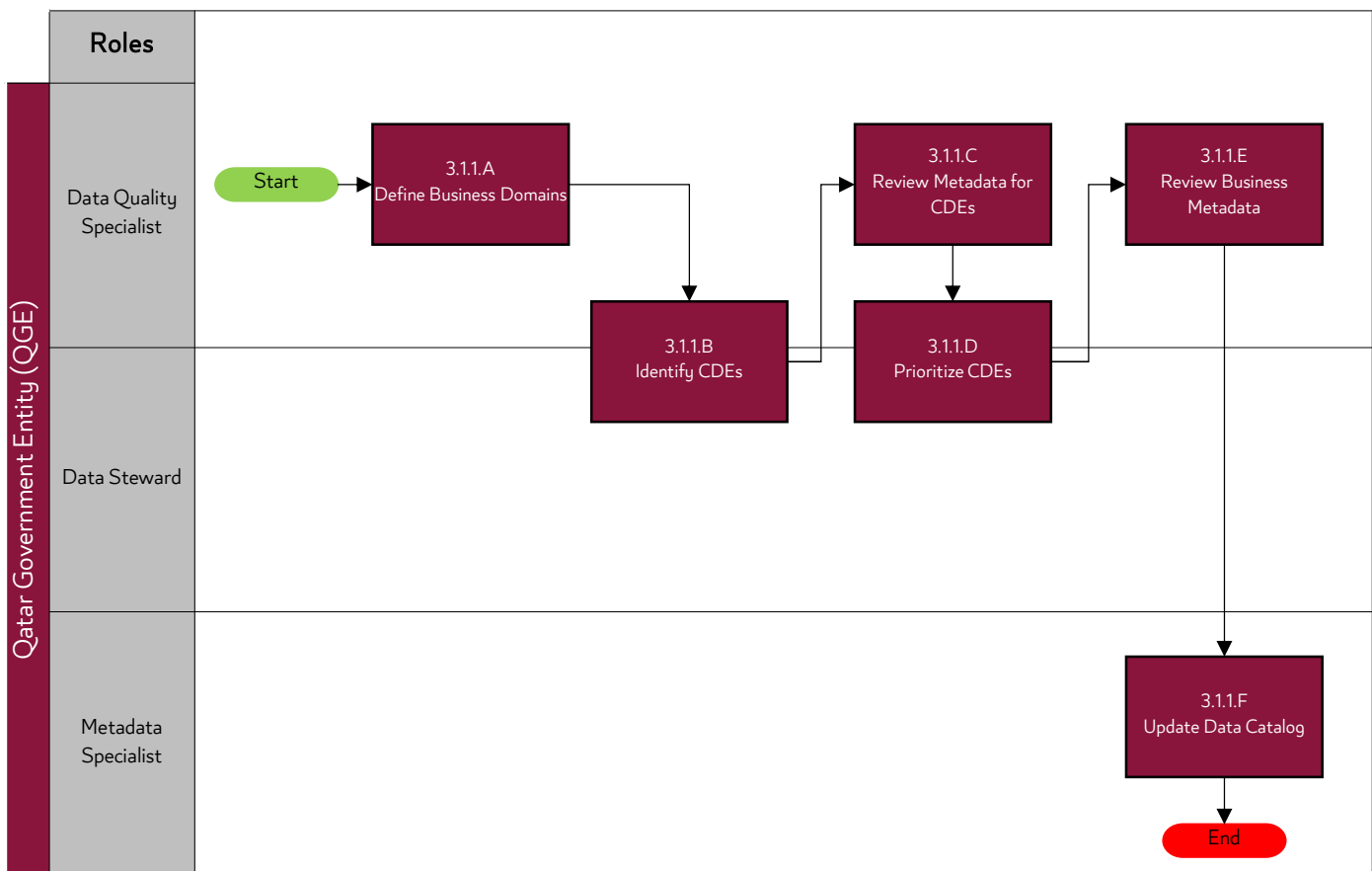


Figure 3: Define Critical Data Elements (CDEs) Process Map

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

## 3.1.1.A

**Define Business Domains:** Identify and document the key business domains that are crucial to the entity's operations. Each domain should be clearly defined based on its operational importance and impact on business objectives. This step ensures that all relevant areas are covered when determining which critical data elements (CDEs) are needed for effective data management and quality control.

## 3.1.1.B

**Identify CDEs:** CDEs are defined by domain data owners as the data elements most widely used across the Entity's business processes, analytics and reporting activities. They are stored across key business applications and represent the data that is critical to the success of the Entity.

**Common criteria for defining CDEs may include:**

- Business impact (e.g., revenue, compliance)
- Frequency of use
- Data quality (e.g., accuracy, completeness)
- Regulatory requirements
- Criticality to business processes

For example, CDEs include data elements used to uniquely identify employees, data elements used to derive values that appear in key management reports, unique identifiers of information important to the business (employee ID, contract ID), and input data elements for use in financial calculations.



**Tip**

While data quality initiatives typically focus on assessing and improving the quality of newly identified CDEs, previously identified CDEs (during Data Strategy phase) should not be ignored. Rather, these should be reviewed in the data catalog and included in the prioritization process.

## 3.1.1.C

**Review Metadata for the CDEs:** The business & technical metadata for the identified CDEs included in the data catalog should be reviewed, this should include the key metadata needed to define the data quality rules.

**Example:****Business Metadata:**

- **Business attribute names:** Employee Nationality
- **Business definition:** represents the nationality country for each employee in the Entity
- **List of values:** List of valid Country Names
- **Reference Data (Y/N):** N
- **Master Data (Y/N):** Y
- **Data Ownership:** Data Owner1
- **Business rules:** Must be a valid country name

**Technical Metadata:**

- **Source System(s):** Organization HRMS
- **Physical data models:** Employee PDM
- **Database:** HR\_DB
- **Schema Name:** HR
- **Table/ View Name:** Employee
- **Attribute Name:** Employee Nationality
- **Primary Key (Y/N):** N
- **Foreign Key (Y/N):** Y
- **Data Type:** String
- **Data Length:** 100 characters
- **Data Format:** DD/MM/YYYY
- **Nullability(Y/N):** Y

**Tip**

In the absence of a data catalog, the above metadata elements should be the starting point for creating and populating the data catalog.

**Key Message / Information**

The data catalog provides the main sources of information to help identify all the data elements that need to be addressed by the data quality initiatives and the respective metadata required to define the data quality rules for these data elements.

## 3.1.1.D

**Prioritize CDEs:** Prioritize CDEs to effectively execute the data quality projects, this can be done based on the following definitions:

- High Criticality: Consistently impacting key activities such as business processes, analytics, or reporting activities
- Medium Criticality: Business process and/or analytics or reporting can be conducted with difficulty
- Low Criticality: Not necessary to conduct business activities or produce analytics or reports

### Asset Guidance

The **Data Quality Critical Data Elements Prioritization Framework and Template** are designed to help Entities prioritize critical data elements (CDEs) for effective data quality management. It provides a structured approach to assess and rank data attributes based on their importance, ensuring focus on the most impactful elements for improved data governance, consistency, and decision-making.

*Illustrative*

|                      | Critical Data Elements Prioritization Framework |
|-------------------------------------------------------------------------------------------------------|-------------------------------------------------|
| The following table reflects the criteria and respective weights to prioritize critical data elements |                                                 |

| 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 |

3.1.1.E

**Review Business Metadata:** Conduct sessions with the data owners to review the business metadata available of the prioritized CDEs in the data catalog to ensure the completeness needed to derive the data quality rules.

**Example:**

- Business terms & definitions for data elements
- The thresholds for the data quality measurement
- Business rules for each data element

**Key Message / Information**

Engaging key stakeholders, including data SMEs and Data Owners, in defining data quality requirements provides a comprehensive understanding of data usage and criticality across business processes and IT systems, and helps capture the necessary metadata for setting data quality rules.

3.1.1.F

**Update Data Catalog:** After reviewing business metadata, update the data catalog to reflect any new changes or additions (e.g., identifying new CDEs). This will ensure that the catalog remains accurate and comprehensive, supporting the creation and maintenance of effective data quality rules and the sustainability of future data initiatives.

## RACI Matrix – Define Critical Data Elements (CDEs)

*Indicative*

| Sr No.  | Activity                 | Data Management and Security Director | Data Management and Governance Section Head | Metadata Specialist | Data Quality Specialist | Business Data Analyst | Data Steward |
|---------|--------------------------|---------------------------------------|---------------------------------------------|---------------------|-------------------------|-----------------------|--------------|
| 3.1.1.A | Define Business Domains  |                                       | A                                           |                     | R                       | C                     | C            |
| 3.1.1.B | Identify CDEs            |                                       | A                                           |                     | R                       | C                     | R            |
| 3.1.1.C | Review Metadata for CDEs |                                       | A                                           |                     | R                       |                       | C            |
| 3.1.1.D | Prioritize CDEs          | I                                     | A                                           |                     | R                       | C                     | R            |
| 3.1.1.E | Review Business Metadata |                                       | A                                           |                     | R                       |                       |              |
| 3.1.1.F | Update Data Catalog      |                                       | A                                           | R                   |                         |                       |              |

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

Based on this Activity, the Entity can



Identify and Prioritize CDEs



Document Relevant Source Systems

### 3.1.2 Conduct Data Profiling

This activity will guide the Entity on the steps to analyze CDEs against the data quality dimensions, identify anomalies, and assess the overall integrity of the data, leading to actionable insights for improvement.

3.1.2.A

**Analyze Data Quality Dimensions:** Assess data quality based on established dimensions in the National Data Quality framework such as accuracy, completeness, consistency, validity and uniqueness. The following are the basic data quality checks that should be carried out on key entities and attributes for each selected source system:

| Data Quality Dimension | Basic Quality Check                       |
|------------------------|-------------------------------------------|
| Accessibility          | Is Data Accessible to Intended Audience?  |
| Accuracy               | Invalid Data Values                       |
| Completeness           | Null Records                              |
| Consistency            | Invalid Format                            |
| Timeliness             | Average Time for Data to Become Available |
| Uniqueness             | Duplicate Records                         |
| Validity               | Invalid Reference Data                    |

3.1.2.B

**Identify Data Anomalies:** Detect any inconsistencies, duplicates, or outliers in the data that may affect its quality. Use data profiling tools to detect irregularities that deviate from expected patterns or rules. For example, look for fields with unexpected values or duplicate entries in a column that should have unique values. Detecting and resolving these anomalies is crucial for maintaining high data quality and ensuring reliable results in downstream processes such as analytics or reporting.

#### RACI Matrix – Conduct Data Profiling

Indicative

| Sr No.  | Activity                        | Data Management and Governance Section Head | Data Quality Specialist | Business Data Analyst | Data Engineer | Data Owner | Data Steward |
|---------|---------------------------------|---------------------------------------------|-------------------------|-----------------------|---------------|------------|--------------|
| 3.1.2.A | Analyze Data Quality Dimensions | A                                           | R                       | R                     |               | R          | C            |
| 3.1.2.B | Identify Data Anomalies         | A                                           | R                       | R                     | C             | R          | R            |

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

Based on this Activity, the Entity can



Profile Previously Identified CDEs

### 3.1.3 Publish Data Profiling Results

This activity will guide the Entity on the steps to create detailed reports that present the data profiling findings clearly to stakeholders, facilitating informed decision-making.

#### 3.1.3.A

**Create Data Profiling Reports:** Summarize findings in detailed reports that highlight key insights, issues, and recommendations. These reports should cover:

- Critical insights
- Highlight any data quality issues (such as missing values or inconsistencies)
- Provide actionable recommendations for addressing these issues
- Include visual aids such as charts or dashboards to make the data more accessible

#### Asset Guidance

The **Data Profiling Report Template** is a report that assesses the quality and characteristics of a dataset. It identifies different metrics and measures of data, helping Entities understand their data better and make informed decisions regarding data management strategies.

*Illustrative*

| Data Profiling Report                                                                                                                                                    |  |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| The following table serves as a template data profiling report to provide a comprehensive analysis of data sets. The Table has been populated with illustrative examples |  |

| Field              | Description                                                                               |
|--------------------|-------------------------------------------------------------------------------------------|
| Data Source ID     | Unique Identifier for the data source                                                     |
| Data Source        | The origin of the data, such as a database, application, or external source.              |
| Table ID           | Unique Identifier for the table                                                           |
| Table Name         | The table in which the data is stored                                                     |
| Attribute ID       | Unique Identifier for the attribute                                                       |
| Attribute Name     | The specific name of the data field being analyzed.                                       |
| Number of Records  | The number of records in the field                                                        |
| Data Type          | The type of data in the field (e.g., integer, string, date).                              |
| Format             | The structure or pattern of the data (e.g., YYYY-MM-DD for dates).                        |
| Accuracy           | The degree to which the data reflects the true values or facts it represents.             |
| Completeness       | The extent to which all required data is present and no values are missing.               |
| Consistency        | The uniformity of data across different datasets or systems, indicating no discrepancies. |
| Timeliness         | The degree to which data is up-to-date and available when needed for decision-making.     |
| Uniqueness         | The measure of how many distinct values are present in a field, indicating no duplicates. |
| Validity           | The extent to which data conforms to defined rules or constraints (e.g., valid formats).  |
| Mean               | The average value of the data in a field, providing a measure of central tendency.        |
| Median             | The middle value in the data set when sorted, representing a measure of central tendency. |
| Mode               | The most frequently occurring value in the data set.                                      |
| Standard Deviation | A measure of the amount of variation or dispersion in the data values.                    |
| Minimum Value      | The smallest value present in the data field.                                             |
| Maximum Value      | The largest value present in the data field.                                              |

| Data Source ID | Data Source Name  | Table ID | Table Name | Attribute ID | Attribute Name | Number of Records | Data Type | Format        |
|----------------|-------------------|----------|------------|--------------|----------------|-------------------|-----------|---------------|
| DS1            | Customer Database | DS1T1    | Customers  | DS1T1A1      | Customer ID    | 1000              | String    | N/A           |
| DS1            | Customer Database | DS1T2    | Sales      | DS1T2A1      | Sale Amount    | 1000              | Float     | Currency(USD) |
| DS1            | Customer Database | DS1T1    | Customers  | DS1T1A2      | Sign-up Date   | 1000              | Date      | YYYY-MM-DD    |
| DS2            | Product Database  | DS2T1    | Products   | DS2T1A1      | Product Name   | 1000              | String    | N/A           |
| DS3            | Product Database  | DS3T1    | Orders     | DS3T1A1      | Order Status   | 1000              | String    | N/A           |

| DQ Dimensions |              |             |            |            |          |
|---------------|--------------|-------------|------------|------------|----------|
| Accuracy      | Completeness | Consistency | Timeliness | Uniqueness | Validity |
| 98%           | 1000         | 1           | 10day      | 1000       | 100%     |
| 95%           | 300          | 1           | 10day      | 320        | 100%     |
| 100%          | 1000         | 1           | 10day      | 730        | 100%     |
| 90%           | 950          | 1           | 10day      | 1000       | 100%     |
| 85%           | 1000         | 1           | 10day      | 300        | 100%     |

| Statistical Summary (Numeric Fields Only) |        |      |                    |               |               |     |
|-------------------------------------------|--------|------|--------------------|---------------|---------------|-----|
| Mean                                      | Median | Mode | Standard Deviation | Minimum Value | Maximum Value |     |
| N/A                                       | N/A    | N/A  | N/A                | N/A           | N/A           | N/A |
| 250.75                                    | 200    | 300  | 75.5               | 50            | 1000          |     |
| N/A                                       | N/A    | N/A  | N/A                | 1/1/2020      | 10/1/2023     |     |
| N/A                                       | N/A    | N/A  | N/A                | N/A           | N/A           | N/A |
| N/A                                       | N/A    | N/A  | N/A                | N/A           | N/A           | N/A |

## 3.1.3.B

**Share with Stakeholders:** Distribute reports to relevant stakeholders for review and action. Ensure that the reports are presented in an easily digestible format, allowing stakeholders to review the key findings and take necessary actions. This step ensures that all parties are informed of any data issues, understand the business impact, and can collaborate effectively to implement the recommendations.

### RACI Matrix – Publish Data Profiling Results

*Indicative*

| Sr No.  | Activity                      | DGO | Data Management and Security Director | Data Management and Governance Section Head | Data Quality Specialist | Business Data Analyst | Data Steward |
|---------|-------------------------------|-----|---------------------------------------|---------------------------------------------|-------------------------|-----------------------|--------------|
| 3.1.5.A | Create Data Profiling Reports |     |                                       | A                                           | R                       | C                     | C            |
| 3.1.5.B | Share with Stakeholders       | I   | I                                     | A                                           | R                       |                       | I            |

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

Based on this Activity, the Entity can



Publish Data Profiling Results

## 3.2 Data Quality Plan

Based on the initial data profiling results, a data quality plan is developed and aligned with the data strategy to guide and monitor the implementation of data quality initiatives across the Entity. The plan should outline key activities and milestones, addressing both quick wins and longer-term strategic improvements. It must also define the roles of owners and stakeholders, highlight interdependencies, assumptions, and risks, and include a detailed implementation timeline and cost structure to ensure successful execution and sustainable data quality management.

### Process – Data Quality Plan

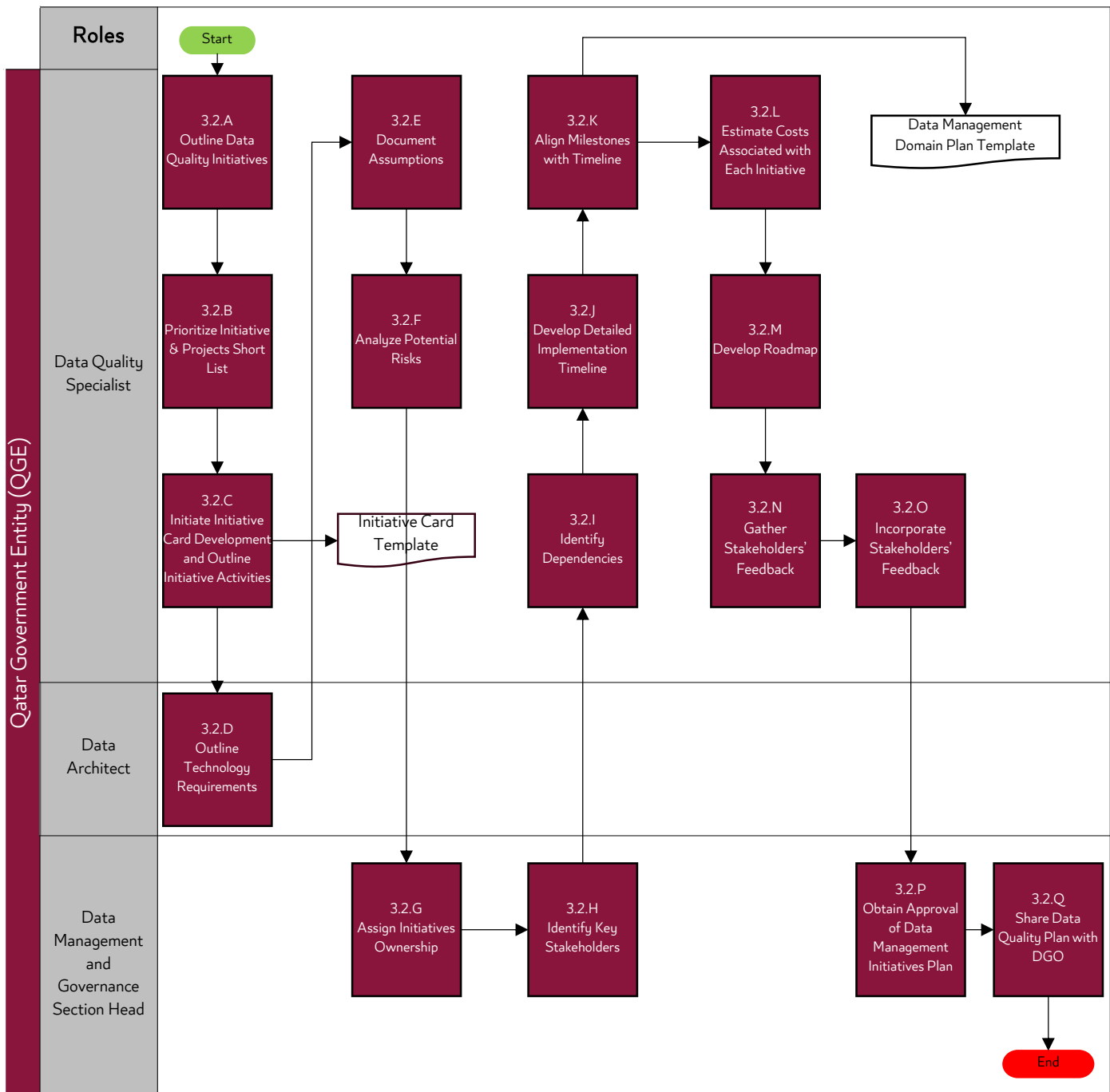


Figure 4: Data Quality Plan Process Map

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

3.2.A

**Outline Data Quality Initiatives:** Identify and document a long list of data quality initiatives, encompassing both short-term and long-term initiatives. Initiatives may include data cleansing, standardization, and validation efforts, as well as larger-scale projects such as system overhauls, governance enhancements, and technology upgrades. Ensure that these initiatives align with the Entity's broader data strategy to support effective resource allocation and address critical data quality needs.

**Key Message / Information**

To ensure alignment with the National Development Strategy (NDS) 3.0, it is critical for the Entity to align its data quality initiatives and projects with the national and sectoral KPIs. Projects should directly contribute to achieving the broader national and sector-specific outcomes outlined in the NDS, supporting strategic priorities and driving measurable impact. By aligning initiatives and projects with sectoral KPIs, the Entity can ensure its efforts are both impactful and consistent with the overarching national vision.

3.2.B

**Prioritize Initiative Short List:** Leverage the data quality initiatives long list to identify and prioritize a short list based on potential impact and feasibility. Focus on high-impact areas where data issues most significantly affect decision-making and operations, ensuring that efforts are directed at enhancing data reliability and supporting strategic initiatives. Additionally, identify quick wins—initiatives that can be implemented rapidly with minimal resources—to provide immediate benefits and demonstrate the value of data quality improvements..

## 3.2.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*

|                                                                                                                                                                                                                                                                             |  |                                |                         |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|--------------------------------|-------------------------|
| <b>OVERVIEW</b> xxx                                                                                                                                                                                                                                                         |  | <b>SPONSOR</b> xxx             | <b>STAKEHOLDERS</b> xxx |
|                                                                                                                                                                                                                                                                             |  | <b>OWNER</b> xxx               |                         |
| <b>KEY ACTIVITIES</b>                                                                                                                                                                                                                                                       |  | <b>TIMELINE</b>                | <b>BUDGET (IN QAR)</b>  |
| 1. Key Activity <ul style="list-style-type: none"> <li>I. Sub-activity X             <ul style="list-style-type: none"> <li>Detailed Steps</li> </ul> </li> <li>II. Sub-activity Y             <ul style="list-style-type: none"> <li>Detailed Steps</li> </ul> </li> </ul> |  | xxx                            | QAR xxx                 |
| 2. Key Activity <ul style="list-style-type: none"> <li>I. Sub-activity X             <ul style="list-style-type: none"> <li>Detailed Steps</li> </ul> </li> <li>II. Sub-activity Y             <ul style="list-style-type: none"> <li>Detailed Steps</li> </ul> </li> </ul> |  | <b>EXPECTED BENEFITS</b>       |                         |
|                                                                                                                                                                                                                                                                             |  | • xxx                          |                         |
|                                                                                                                                                                                                                                                                             |  | <b>KEY MILESTONES</b>          |                         |
|                                                                                                                                                                                                                                                                             |  | A. xxx<br>B. xxx<br>C. xxx     |                         |
|                                                                                                                                                                                                                                                                             |  | <b>TECHNOLOGY REQUIREMENTS</b> |                         |
|                                                                                                                                                                                                                                                                             |  | • xxx                          |                         |

|                       |                             |                                |                                        |
|-----------------------|-----------------------------|--------------------------------|----------------------------------------|
| <b>OVERVIEW</b> xxx   |                             | <b>SPONSOR</b> xxx             | <b>STAKEHOLDERS</b> xxx                |
|                       |                             | <b>OWNER</b> xxx               |                                        |
| <b>KEY ACTIVITIES</b> |                             | <b>TIMELINE</b>                |                                        |
| #                     | 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.                             |                                        |
| <b>DEPENDENCIES</b>   |                             | <b>ASSUMPTIONS &amp; RISKS</b> |                                        |
| • xxx                 |                             | • xxx                          |                                        |

3.2.D

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

3.2.E

**Document Assumptions:** Record assumptions regarding resource availability, stakeholder engagement, and data complexity to ensure transparency and guide decision-making.

3.2.F

**Analyze Potential Risks:** Clearly identify and assess the risks associated with data quality initiatives. This includes:

- Technical risks such as system downtime, data corruption, or tool compatibility issues.
- Operational risks such as delays due to resource shortages or conflicting priorities between teams
- Financial risks such as budget overruns due to unforeseen complexity or extended timelines

3.2.G

**Assign Initiatives Ownership:** The Data Management and Governance Section Head & Data Architecture and Analytics Section Head will be the owner of all data quality initiatives. This designation ensures that the initiatives have a dedicated leader who can oversee progress and resolve roadblocks that may arise.



Tip

Be sure to appoint clear initiative owners, this will drive overall accountability and operational focus.

3.2.H

**Identify Key Stakeholders:** Recognize individuals and groups that have a vested interest in data quality, including data users, IT staff, and management, facilitating collaboration and support. Understanding who needs to be engaged helps streamline communication and decision-making, improving the overall effectiveness of the data quality program.

3.2.I

**Identify Dependencies:** Analyze how Data Quality initiatives interact with other projects and processes to streamline resource allocation and improve coordination. Key considerations include:

- **Reviewing Impacts on Organizational Objectives:** Evaluate how Data Quality initiatives influence or align with overarching business goals and strategic initiatives
- **Identifying Overlaps in Resources and Teams:** Assess where the same personnel, tools, or systems are involved in multiple initiatives to avoid resource conflicts
- **Ensuring Cross-Departmental Collaboration:** Coordinate between departments to resolve competing priorities and optimize shared resources for Data Quality projects.

3.2.J

**Develop Detailed Implementation Timeline:** Break down each activity in an initiative into actionable steps with start and end dates, ensuring the timeline is realistic and achievable. Each step should have a clear description of the required task, dependencies on other activities, and allocated resources. Ensure that the timeline accounts for potential delays or risks identified earlier, so adjustments can be made without derailing the project.

3.2.K

**Align Milestones with Timeline:** Set specific checkpoints or milestones throughout the implementation to assess progress. These milestones allow the evaluation of whether the initiative is proceeding as planned and provide opportunities for course corrections.

### Asset Guidance

The **Data Management Domain Plan Template** helps you capture an overview of each identified initiative, detail its sub-activities, owners, stakeholders, and activity timeframe

*Illustrative*

| Data Management Initiative Plan Template                                                        |  |  |  |  |  |  |  |  |  |
|-------------------------------------------------------------------------------------------------|--|--|--|--|--|--|--|--|--|
| The following table is a template capturing a brief overview of the domain relevant initiatives |  |  |  |  |  |  |  |  |  |

| Field                  | Description                                                                                           |
|------------------------|-------------------------------------------------------------------------------------------------------|
| Initiative ID          | A unique identifier for each initiative to facilitate tracking and reference                          |
| Initiative Title       | the name of the initiative                                                                            |
| Initiative Description | A short description of the initiative, outlining its objectives and expected outcomes                 |
| Cost                   | Estimated financial resources required for the initiative, including both direct and indirect costs   |
| Start Date             | The date when the initiative is scheduled to begin                                                    |
| End Date               | The date by which the initiative is expected to be completed                                          |
| Interdependencies      | Any other identified initiatives or projects that this initiative depends on for successful execution |
| Assumptions            | Conditions assumed to be true for the initiative's success (e.g., resource availability)              |
| Risks                  | Potential risks that could affect the initiative's success                                            |

| Initiative ID | Initiative Title | Initiative Description | Cost | Start Date | End Date | Interdependencies | Assumptions | Risks |
|---------------|------------------|------------------------|------|------------|----------|-------------------|-------------|-------|
|               |                  |                        |      |            |          |                   |             |       |
|               |                  |                        |      |            |          |                   |             |       |
|               |                  |                        |      |            |          |                   |             |       |
|               |                  |                        |      |            |          |                   |             |       |
|               |                  |                        |      |            |          |                   |             |       |

| Initiative ID | Initiative Title | Sub-activities | Owner | Stakeholder(s) | Status      | Activity Timeframes |          | Roadmap |       |        |     |
|---------------|------------------|----------------|-------|----------------|-------------|---------------------|----------|---------|-------|--------|-----|
|               |                  |                |       |                |             | Start Date          | Due Date | 0       | 1     | 2      | 3   |
|               |                  |                |       |                |             |                     |          | 1 Jan   | 8 Jan | 15 Jan | ### |
| DMISG-001     |                  |                |       |                | In Progress | 1 Jul               | 30 Dec   |         |       |        |     |
| DMISG-002     |                  |                |       |                | Completed   | 1 Jul               | 31 Dec   |         |       |        |     |
| DMISG-003     |                  |                |       |                | Not Started | 30 Mar              | 31 Jul   |         |       |        |     |

3.2.L

***Estimate Costs Associated with Each Initiative:*** Consider expenses related to tools, technology, personnel, and training required for effective implementation. Break down costs based on the needs of each task, such as acquiring software for data validation, hiring additional staff, or providing necessary training for existing teams. This step ensures the budget is well-planned, realistic, and includes all critical aspects to enable smooth implementation of the initiatives.

3.2.M

***Develop Roadmap:*** Create an overall roadmap that outlines timelines, phases, and key milestones for implementing the prioritized data quality initiatives. Sequence the initiatives based on dependencies, resource availability, and alignment with organizational goals. The roadmap should also include projected timelines for each initiative, allowing stakeholders to monitor progress and allocate resources effectively over time.

3.2.N

***Gather Stakeholders' Feedback:*** Present the draft plan to all relevant parties, inviting input and suggestions to enhance its effectiveness. This feedback process helps surface any potential issues or overlooked areas, improving the plan's effectiveness and ensuring all voices are heard. Engaging stakeholders early also builds buy-in and commitment to the plan's success.

3.2.O

***Incorporate Stakeholders' Feedback:*** Actively gather and incorporate constructive feedback from key stakeholders to refine the plan and ensure it aligns with their needs and expectations. This collaborative approach helps to address concerns early, fine-tune the details, and strengthen the overall effectiveness of the plan.

3.2.P

***Obtain Approval of Data Quality Plan:*** After incorporating all feedback, the Data Management and Security Director will be presented with the finalized Data Quality Plan for review and approval. This ensures alignment with the Entity's priorities and secures the necessary support for successful implementation. Once approved, the plan can be formally adopted and published.

3.2.Q

***Share Data Quality Plan with DGO:*** Submit the finalized Data Quality 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 – Data Quality Plan

*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 Quality Specialist | Data Strategy Expert | Data Steward | Data Architect |
|--------|------------------------------------------------------------------------|-----|---------------------------------------|---------------------------------------------|----------------------------------------------|----------------------------|-------------------------|----------------------|--------------|----------------|
| 3.2.A  | Outline Data Quality Initiatives                                       |     | I                                     | A                                           |                                              | C                          | R                       | C                    | C            |                |
| 3.2.B  | Prioritize Initiative & Projects Short List                            |     | I                                     | A                                           |                                              | C                          | R                       | C                    | C            |                |
| 3.2.C  | Initiate Initiative Card Development and Outline Initiative Activities |     | I                                     | A                                           |                                              | C                          | R                       | C                    | C            |                |
| 3.2.D  | Outline Technology Requirements                                        |     |                                       | A                                           | C                                            |                            | C                       | C                    |              | R              |
| 3.2.E  | Document Assumptions                                                   |     |                                       | A                                           | C                                            | C                          | R                       | C                    | C            | C              |
| 3.2.F  | Analyze Potential Risks                                                |     |                                       | A                                           | C                                            | C                          | R                       | C                    | C            | C              |
| 3.2.G  | Assign Initiatives Ownership                                           |     | A                                     | R                                           | C                                            |                            |                         |                      |              |                |
| 3.2.H  | Identify Key Stakeholders                                              |     | A                                     | R                                           | C                                            |                            |                         |                      |              |                |
| 3.2.I  | Identify Dependencies                                                  |     | I                                     | A/C                                         | C                                            |                            | R                       | C                    | C            |                |
| 3.2.J  | Develop Detailed Implementation Timeline                               |     | I                                     | A                                           | C                                            | C                          | R                       | C                    | C            |                |
| 3.2.K  | Align Milestones with Timeline                                         |     | I                                     | A                                           | C                                            | C                          | R                       | C                    | C            |                |
| 3.2.L  | Estimate Costs Associated with Each Initiative                         |     | I                                     | A/C                                         | C                                            |                            | R                       | C                    | C            |                |
| 3.2.M  | Develop Roadmap                                                        |     | I                                     | A/C                                         | C                                            |                            | R                       | C                    | C            |                |
| 3.2.N  | Gather Stakeholders' Feedback                                          |     | I                                     | A                                           |                                              | C                          | R                       |                      | C            |                |
| 3.2.O  | Incorporate Stakeholders' Feedback                                     |     | I                                     | A                                           | C                                            | C                          | R                       | C                    | C            |                |
| 3.2.P  | Obtain Approval of Data Management Initiatives Plan                    |     | A                                     | R                                           |                                              |                            |                         |                      |              |                |
| 3.2.Q  | Share Data Quality 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 Data Quality Initiatives



Estimate Timelines and Costs



Develop and Review Data Quality Plan

## 4. Data Quality Implementation



## 4. Data Quality Implementation

### What is Data Quality Implementation and why does the Entity need it?

The Data quality implementation phase translates strategic planning into actionable practices, ensuring that high standards of data quality are upheld across the Entity. This phase involves establishing specific data quality rules and thresholds that define acceptable quality levels for key data elements, as well as deploying tools that automate monitoring and reporting processes. Effective implementation is essential because it empowers Entities to consistently identify and rectify data quality issues, enhance operational efficiency, and ensure compliance with regulatory requirements. By leveraging robust data quality tools and clearly defined standards, Entities can foster a culture of data excellence, enabling informed decision-making and driving overall business success.

#### Data Quality Implementation Key Benefits:

- **Ensures Accurate and Reliable Reporting:** Implementing data quality management ensures data accuracy and consistency, providing a reliable foundation for generating precise, error-free reports that support informed decision-making and build stakeholder confidence.
- **Establishes Consistent Data Standards:** By setting clear data quality metrics, Entities create a uniform standard for evaluating data quality, ensuring that all teams adhere to the same expectations and practices across the Entity.
- **Increases Operational Efficiency:** The deployment of data quality tools automates monitoring and reporting processes, allowing teams to focus on strategic activities rather than manual checks, ultimately improving productivity and reducing resource waste.
- **Strengthens Compliance and Governance:** Effective data quality implementation ensures that data adheres to regulatory requirements and internal policies, minimizing the risk of non-compliance and enhancing data governance efforts.
- **Fosters Collaboration Across Departments:** With clearly defined roles and responsibilities in data quality management, teams are better positioned to collaborate, share insights, and align efforts towards common data quality goals.
- **Promotes Data-Driven Culture:** By prioritizing data quality, Entities cultivate a culture that values accurate and reliable data, empowering teams to make informed decisions that drive better strategic outcomes and enhance competitiveness.

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



## Guiding Principles for Data Quality Implementation

1. **Establish Clear Data Quality Rules and Thresholds:** Develop data quality rules and thresholds that are aligned with the Entity's objectives, ensuring consistency and clarity in what constitutes acceptable data quality.
2. **Engage Stakeholders Early:** Involve stakeholders from various departments early in the implementation process to gather diverse perspectives and foster a collective ownership of data quality initiatives.
3. **Implement Robust Monitoring Mechanisms:** Establish systems to track data quality in real time, allowing for timely detection of issues and flexibility to adapt strategies as business needs and technologies evolve.
4. **Provide Training and Resources:** Develop training programs for teams involved in data quality and equip them with the necessary tools and resources to effectively uphold data quality standards.
5. **Embed Quality Practices:** Integrate data quality checks and processes into existing business workflows to ensure data quality is a fundamental part of operations.

## 4.1 Data Quality Rules & Thresholds

Data quality rules define the standards for data attributes, specifying conditions like accuracy, completeness, and consistency for key data fields. Thresholds set limits on these standards, such as the maximum acceptable error rate. By documenting these rules and thresholds, Entities ensure consistent data validation, enabling early identification of issues and clear remediation actions like rejection or correction. This process is crucial for maintaining reliable, high-quality data that supports effective decision-making and aligns with both business and technical requirements.

### 4.1.1 Define Data Quality Rules

This activity will guide the Entity on the steps to define technical and business rules to specify the criteria for assessing a data quality dimension.

#### 4.1.1.A

**Define Data Quality Rules:** Develop clear data quality rules for Critical Data Elements (CDEs) in collaboration with business and technical stakeholders. Each rule should include a detailed description, type (business or technical), applicable data fields, condition and logic, relevant data quality dimension, thresholds, remediation actions, rule owner, and data steward. This ensures data quality standards are well-defined, measurable, and actionable.

**Technical Data Quality Rules:** validate the integrity of the data and can be drawn from technical metadata within the data catalog. It includes rules such as:

- Primary Key enforcement
- Foreign Key existence
- Nullability
- Data Format

**Business Data Quality Rules:** are defined in alignment with business scenarios and the functional requirements of the data. They are derived from the business rules defined in the business metadata of the data catalog. Examples of business rules:

- QID expiry date should be in the future
- Nationality should have a valid country name

4.1.1.A  
Cont. Asset Guidance

A **Data Quality Rules Template** is used to define, document, and manage criteria for assessing the quality of data within an Entity. It typically includes components such as a unique rule ID, a descriptive rule name, a detailed description of the rule's purpose, the specific data domain and elements it applies to, the type of data quality check (e.g., validity, completeness, accuracy), and the logical expression or formula defining the rule.

*Illustrative*

| Data Quality Rules                                                                                                                                    |                                                                                                                                                                              |  |  |  |  |  |  |  |  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|--|--|--|--|--|--|--|
| The following table is a template containing the business and technical rules for data quality. The table has been populated with illustrative values |                                                                                                                                                                              |  |  |  |  |  |  |  |  |
| Field                                                                                                                                                 | Description                                                                                                                                                                  |  |  |  |  |  |  |  |  |
| Data Source ID                                                                                                                                        | Unique Identifier for the data source                                                                                                                                        |  |  |  |  |  |  |  |  |
| Data Source                                                                                                                                           | The origin of the data, such as a database, application, or external source.                                                                                                 |  |  |  |  |  |  |  |  |
| Table ID                                                                                                                                              | Unique Identifier for the table                                                                                                                                              |  |  |  |  |  |  |  |  |
| Table Name                                                                                                                                            | The table in which the data is stored                                                                                                                                        |  |  |  |  |  |  |  |  |
| Attribute ID                                                                                                                                          | Unique Identifier for the attribute                                                                                                                                          |  |  |  |  |  |  |  |  |
| Attribute Name                                                                                                                                        | The specific name of the data field to which the rule applies                                                                                                                |  |  |  |  |  |  |  |  |
| Rule ID                                                                                                                                               | Unique identifier for the data quality rule                                                                                                                                  |  |  |  |  |  |  |  |  |
| Rule Description                                                                                                                                      | A detailed explanation of the data quality rule, including its purpose and importance.                                                                                       |  |  |  |  |  |  |  |  |
| Rule Type                                                                                                                                             | Indicates whether the rule is a business rule (related to business processes) or a technical rule (related to system design and implementation).                             |  |  |  |  |  |  |  |  |
| Rule Condition and Business Logic                                                                                                                     | Defines the specific conditions or criteria that must be met for the rule to be triggered, outlining the associated logic.                                                   |  |  |  |  |  |  |  |  |
| Relevant Data Quality Dimension                                                                                                                       | Identifies the specific data quality dimension (e.g., accuracy, completeness) that the rule addresses, as per the data quality management framework.                         |  |  |  |  |  |  |  |  |
| Data Quality Thresholds                                                                                                                               | Sets the acceptable limits or ranges for the data quality metrics, indicating what levels of quality are considered satisfactory.                                            |  |  |  |  |  |  |  |  |
| Recommended Remedial Action                                                                                                                           | Suggests actions to take when the rule is violated, such as rejection, correction, or notification of stakeholders.                                                          |  |  |  |  |  |  |  |  |
| Rule Owner                                                                                                                                            | Designates the individual or role responsible for the rule, typically a business user, business analyst, or subject matter expert.                                           |  |  |  |  |  |  |  |  |
| Data Steward                                                                                                                                          | Identifies the person or role tasked with managing and overseeing the data quality rule, often an IT professional, data architect, data engineer, or database administrator. |  |  |  |  |  |  |  |  |
| Approved By                                                                                                                                           | Name of the data quality rule approver                                                                                                                                       |  |  |  |  |  |  |  |  |
| Date of Approval                                                                                                                                      | Date for when the data quality rule was approved                                                                                                                             |  |  |  |  |  |  |  |  |

| Data Source ID | Data Source       | Table ID | Table Name   | Attribute ID | Attribute Name   | Rule ID | Rule Description                                       | Rule Type      | Rule Condition and Business Logic                                      |
|----------------|-------------------|----------|--------------|--------------|------------------|---------|--------------------------------------------------------|----------------|------------------------------------------------------------------------|
| DS1            | Customer Database | DS1T1    | Customers    | DS1TA1       | Customer Email   | DQR1    | Ensure all customer email addresses are valid formats. | Business Rule  | Email must match the regex pattern for valid email addresses.          |
| DS1            | Customer Database | DS1T1    | Customers    | DS1TA2       | First Name       | DQR2    | All required fields must be completed for new entries. | Business Rule  | All fields must not be null for new customer records.                  |
| DS1            | Customer Database | DS1T1    | Customers    | DS1TA3       | Customer ID      | DQR3    | Customer IDs must be unique across the database.       | Technical Rule | No duplicates allowed in the Customer ID field.                        |
| DS2            | Sales Database    | DS2T1    | Transactions | DS2TA1       | Transaction Date | DQR4    | Transaction dates must be within the last 30 days.     | Business Rule  | Transaction Date must be greater than or equal to today minus 30 days. |
| DS3            | Product Database  | DS3T1    | Products     | DS3TA1       | Product Price    | DQR5    | Product prices must be positive values.                | Technical Rule | Price must be greater than zero.                                       |

| Relevant Data Quality Dimension | Data Quality Thresholds | Recommended Remediation Action         | Rule Owner            | Data Steward           | Approved By | Date of Approval |
|---------------------------------|-------------------------|----------------------------------------|-----------------------|------------------------|-------------|------------------|
| Validity                        | 100% valid              | Notification to correct invalid emails | Business Analyst      | Data Engineer          | Person 1    | 10/3/2024        |
| Completeness                    | 100% filled             | Rejection of incomplete entries        | Business User         | Data Architect         | Person 1    | 10/3/2024        |
| Uniqueness                      | 0% duplicates           | Correction of duplicate entries        | Subject Matter Expert | Database Administrator | Person 1    | 10/3/2024        |
| Timeliness                      | 100% timely             | Notification for outdated entries      | Business Analyst      | Data Steward           | Person 1    | 10/6/2024        |
| Validity                        | > 0                     | Correction of negative prices          | Business User         | Data Engineer          | Person 1    | 12/8/2024        |

## RACI Matrix – Define Data Quality Rules

*Indicative*

| Sr No.  | Activity                  | Data Management and Governance Section Head | Data Owner | Data Steward | Data Quality Specialist | Business Data Analyst |
|---------|---------------------------|---------------------------------------------|------------|--------------|-------------------------|-----------------------|
| 4.1.1.A | Define Data Quality Rules | A                                           | C          | R            | R                       | C                     |

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

Based on this Activity, the Entity can



Define Technical & Business Data Quality Rules

### 4.1.2 Define Data Quality Metrics and Thresholds

This activity will guide the Entity in defining data quality metrics to be measured for each data quality dimension, and the respective thresholds for each metric.

#### 4.1.2.A

**Define Data Quality Metrics:** For each data quality dimension, define the data quality metric, which is the unit of measurement for the quality of data based on the rules defined for each dimension.

**Example:**

- **Completeness:** % of records with non-empty values for a mandatory data attribute
- **Validity:** % of records compliant with the defined validity dimension data quality rules

#### 4.1.2.B

**Define Data Quality Thresholds:** For each data quality metric define the acceptance threshold which is a % or number of records that defines the acceptance level of a data quality metric. Data quality thresholds should be defined based on the criticality of the data element being assessed.

**Example:**

100% of the records in the EMPLOYEE table must have First Name and Last Name values, else, this will be flagged as an issue and escalated accordingly.



Tip

Not all thresholds should be set to 100%, as there needs to be a balance between the impact of not having a 100% Data Quality threshold and the cost and feasibility of having a 100% Data Quality threshold for a specific Data Quality rule.

### RACI Matrix – Define Data Quality Metrics and Thresholds

*Indicative*

| Sr No.  | Activity                       | Data Management and Governance Section Head | Data Owner | Data Steward | Data Quality Specialist | Business Data Analyst |
|---------|--------------------------------|---------------------------------------------|------------|--------------|-------------------------|-----------------------|
| 4.1.2.A | Define Data Quality Metrics    | A                                           | C          | R            | R                       | C                     |
| 4.1.2.B | Define Data Quality Thresholds | A                                           | C          | R            | R                       | C                     |

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

Based on this Activity, the Entity can



Define Metrics & Thresholds for Data Quality Dimensions

## 4.2 Data Quality Tool Implementation

Implementing a data quality tool helps automate and streamline the management of data quality across an Entity. It enables continuous monitoring of data by automatically profiling it to detect anomalies and inconsistencies. The tool also facilitates the creation and execution of data quality rules to assess data against predefined standards, generating reports on any identified issues. This ensures that data remains accurate, complete, and consistent. Advanced features may include data cleansing, integration with external systems, issue tracking, and collaboration capabilities for data stakeholders. Implementing such a tool improves data reliability, supports proactive data management, and enhances decision-making processes.

### 4.2.1 Establish Functional and Technical Requirements

This activity will guide the Entity on the steps to identify and prioritize the key features and functionalities needed in a data quality tool.

4.2.1.A

**Identify Key Features:** Determine the essential features that the data quality monitoring tool must possess to meet the Entity's data quality objectives such as data profiling, cleansing, standardization, etc. These features should align with the business and technical requirements.

4.2.1.B

**Prioritize Tool Requirements:** Prioritize features based on their importance and impact on data quality goals, distinguishing between must-have and nice-to-have functionalities. This can be achieved by assigning weights to each feature of the data quality monitoring tool according to its significance and criticality, which helps the Entity make informed decisions during selection, ensuring essential functionalities are emphasized and the tool aligns with strategic goals for data quality improvement.

### RACI Matrix – Establish Functional and Technical Requirements

Indicative

| Sr No.  | Activity                     | DGO | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Quality Specialist | Business Data Analyst | Data Architect | Data Strategy Expert |
|---------|------------------------------|-----|---------------------------------------|---------------------------------------------|----------------------------|-------------------------|-----------------------|----------------|----------------------|
| 4.2.1.A | Identify Key Features        |     | I                                     | A                                           | R                          | R                       | C                     | C              |                      |
| 4.2.1.B | Prioritize Tool Requirements | I   | I                                     | A                                           | R                          | R                       | C                     | C              | I                    |

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

Based on this Activity, the Entity can



Define Data Quality Tool Requirements

### 4.2.2 Conduct Tool Evaluation and Select Suitable Tool

This activity will guide the Entity on the steps to research potential data quality tools and plan the evaluation process for selecting the most suitable option.

#### Process - Conduct Tool Evaluation and Select Suitable Tool

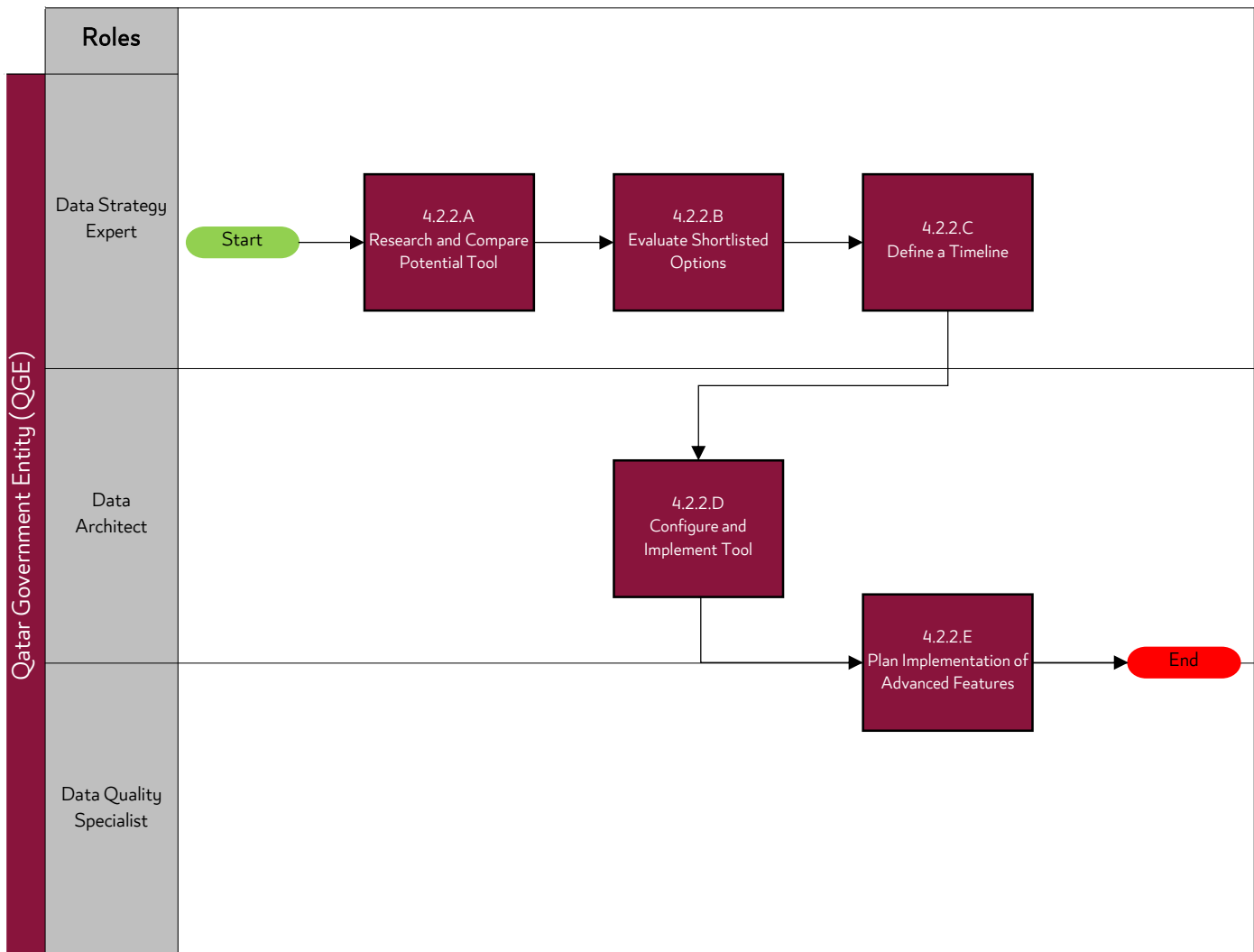


Figure 5: Conduct Tool Evaluation and Select Suitable Tool Process Map

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

#### 4.2.2.A

**Research and Compare Potential Tools:** Begin by identifying the specific requirements the Entity needs in a data quality tool, such as data cleansing, profiling, and monitoring capabilities. Compile a list of potential options by utilizing industry reports and user feedback while considering vendor reputation and support. Gather pricing information, evaluating both initial costs and long-term expenses to identify the most cost-effective solutions. Finally, assess each tool's compatibility with the existing systems to shortlist potential tools.

## 4.2.2.B

**Evaluate Shortlisted Options:** After shortlisting a few tools, the evaluation process can begin. The evaluation criteria have different weights based on feature importance to business processes and are outlined in the Data Quality Tool Selection Criteria Template. Additionally, the evaluation process may include demos and trials to ensure selection of the best option.

### Asset Guidance

The **Data Quality Tool Selection Criteria Template** is a structured framework to help Entities evaluate and choose suitable data quality tools. It includes essential criteria for assessment, such as functional requirements, technical specifications, support options, and budget considerations.

*Illustrative*

|  <b>Data Quality Management Tool Selection Criteria Template</b> |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------|--|
| The following tables provides a structured template to guide the evaluation and selection of a data quality 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<br>[Tool name 1,2,3,X] | The level of advancement of the feature as per the standards (Foundational, Advanced, or Very Advanced).<br>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                                         |                     |                              |               |               |               |               |
| Automated Data Profiling                     |                     |                              |               |               |               |               |
| Creating Data Quality Rules                  |                     |                              |               |               |               |               |
| Conducting Data Quality Assessments          |                     |                              |               |               |               |               |
| Generating Data Quality Assessment Reports   |                     |                              |               |               |               |               |
| Automated Data Cleansing and Standardization |                     |                              |               |               |               |               |
| API Integration Support                      |                     |                              |               |               |               |               |
| Data Quality Issue Logging                   |                     |                              |               |               |               |               |
| Collaboration Features                       |                     |                              |               |               |               |               |
| AI-Powered Data Quality Management           |                     |                              |               |               |               |               |
| Total                                        |                     |                              | 0%            | 0%            | 0%            | 0%            |

## 4.2.2.C

**Define a Timeline:** Establish a detailed timeline for the implementation of the data quality monitoring tool. This timeline will outline key phases, milestones, and deadlines, ensuring that the project is completed efficiently and on schedule. A well-defined timeline helps coordinate efforts among stakeholders, manage expectations, and facilitate effective project management.

## 4.2.2.D

**Configure and Implement Tool:** Upon selecting the most suitable option, set up the tool, customize its features, integrate it with existing systems, and implement necessary configurations to ensure it operates effectively within the Entity's data environment. Proper configuration and implementation are crucial for maximizing the tool's effectiveness and ensuring it meets data quality objectives.

## 4.2.2.E

**Plan Implementation of Advanced Features:** After successfully implementing the tool with the fundamental and essential functionalities, set a timeline and plan for implementation of advanced, nice-to-have features in the future to support the long-term strategic objectives of data quality management.

### RACI Matrix – Conduct Tool Evaluation and Select Suitable Tool

*Indicative*

| Sr No.  | Activity                                 | DGO | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Quality Specialist | Data Architect | Data Strategy Expert |
|---------|------------------------------------------|-----|---------------------------------------|---------------------------------------------|----------------------------|-------------------------|----------------|----------------------|
| 4.2.2.A | Research and Compare Potential Tool      |     | I                                     | A                                           |                            | C                       | C              | R                    |
| 4.2.2.B | Evaluate Shortlisted Options             | I   | I                                     | A                                           | C                          | C                       | C              | R                    |
| 4.2.2.C | Define a Timeline                        | I   | I                                     | A                                           | C                          | C                       | C              | R                    |
| 4.2.2.D | Configure and Implement Tool             |     |                                       | A                                           |                            | C                       | R              |                      |
| 4.2.2.E | Plan Implementation of Advanced Features | I   | I                                     | A                                           |                            | R                       | R              |                      |

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

Based on this Activity, the Entity can



Evaluate and Select Suitable Data Quality Tool



Monitor Data Quality Metrics

### 4.2.3 Conduct Data Quality Assessment

This activity will guide the Entity on the steps to conduct the Data Quality Assessment after selecting and implementing its data quality tool.

#### 4.2.3.A

**Conduct Data Quality Assessment.** Apply predefined data quality rules to the Critical Data Elements (CDEs) to evaluate their current quality levels. This assessment will identify areas for improvement and provide a clear picture of data health. The results should be integrated into the Data Quality Dashboard (*covered in 6.2.A*) to generate detailed reports, offering actionable insights and guiding data remediation efforts.

The Data Quality Assessment shall be conducted in accordance with the National Data Quality Framework, which is based on seven dimensions derived from the National Quality Assurance Framework (NQAF) as follows:

1. **Accuracy:** Data accurately mirrors real-world objects or events, aligning with trusted sources.
2. **Completeness:** All essential data is available to meet both current and future needs.
3. **Consistency:** Data from different sources aligns without conflicts across records or attributes.
4. **Uniqueness:** Each record is unique, appearing only once and easily identifiable across systems.
5. **Timeliness:** Data is delivered exactly when needed, ensuring relevance.
6. **Validity:** Data adheres to required formats and business rules.
7. **Accessibility:** Data is readily available, easy to retrieve, and use across platforms.

#### Asset Guidance

The **Data Quality Assessment Template** is designed to provide entities with a structured approach to evaluating the quality of their datasets across multiple dimensions, as defined by the National Data Quality Framework.

*Illustrative*

| Dimension                                                                     | Sub-dimension                                | Description                                                                                                                                   | Scoring Criteria                                                                                                                                                                                  | Score for Dataset #1 | Score for Dataset #2 | Score for Dataset #3 | Score for Dataset #4 |
|-------------------------------------------------------------------------------|----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|----------------------|----------------------|----------------------|
| <b>Example:</b>                                                               | Semantic Accuracy                            | Measures whether data values make logical sense in context and adhere to real-world knowledge.                                                | Significant issues, more than 50% of values are unrealistic.                                                                                                                                      | 1                    | 1                    | 1                    | 1                    |
| <b>Accuracy</b><br><i>Is the data accurate and to what level?</i>             | Semantic Accuracy                            | Measures whether data values make logical sense in context and adhere to real-world knowledge.                                                | Score 3: All values fall within realistic ranges with no logically inconsistent data.<br>Score 2: Minor unrealistic values; up to 50% of data falls outside expected ranges.                      | 1                    | 2                    | 1                    | 1                    |
|                                                                               | Precision                                    | Measures the level of detail and granularity in data values, ensuring that they are as specific as required for the purpose.                  | Score 3: Data is collected and stored with the required granularity.<br>Score 2: Minor rounding issues; up to 50% of data is less precise than required.<br>Score 1: Major precision issues; more | 1                    | 2                    | 2                    | 2                    |
| <b>Completeness</b><br><i>How complete is the data? Are there known gaps?</i> | Field Completeness (Not Null Fields)         | Assesses the degree to which fields that must not allow null values are fully populated with non-null data                                    | Score 3: All mandatory non-null fields are fully populated.<br>Score 2: Up to 50% of mandatory fields are null.<br>Score 1: More than 50% of mandatory non-null fields are null.                  | 1                    | 1                    | 2                    | 2                    |
|                                                                               | Field Completeness (Null Fields)             | Measures the degree to which fields that do not have a specific requirement on the presence of non-null values are filled with non-null data. | Score 3: All non-mandatory fields are fully populated.<br>Score 2: Up to 50% of non-mandatory fields are null.<br>Score 1: More than 50% of non-mandatory fields are null.                        | 3                    | 3                    | 2                    | 2                    |
|                                                                               | Metadata Completeness (Mandatory Attributes) | The extent to which metadata describing mandatory attributes (e.g., definitions, formats, lineage) is fully populated and documented.         | Score 3: Metadata for all mandatory attributes is complete and accurate.<br>Score 2: Up to 50% of metadata for mandatory attributes is missing or incomplete.                                     | 2                    | 2                    | 3                    | 3                    |

## 4.2.3.B

**Post-Assessment Review and Action Plan:** After completing the assessment, review the results with key stakeholders, including Data Stewards, to interpret findings, identify areas requiring remediation, and notify the DGO upon completion of the assessment.

### RACI Matrix – Conduct Data Quality Assessment

*Indicative*

| Sr No.  | Activity                               | DGO | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Quality Specialist | Data Steward |
|---------|----------------------------------------|-----|---------------------------------------|---------------------------------------------|----------------------------|-------------------------|--------------|
| 4.2.3.A | Conduct Data Quality Assessment        |     | I                                     | A                                           |                            | R                       |              |
| 4.2.3.B | Post-Assessment Review and Action Plan | I   |                                       | A                                           | C                          | R                       | C            |

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

Based on this Activity, the Entity can



Conduct Data Quality Assessment

## 5. Data Quality Issue Management



## 5. Data Quality Issue Management

### What is Data Quality Issue Management and why does the Entity need it?

Data quality issue management is a process that promotes the continuous maintenance of high data quality standards by identifying, tracking, and resolving data quality issues across the Entity. This process involves the systematic identification of data discrepancies and errors, followed by thorough root cause analysis to understand the underlying problems. It includes the creation of a structured workflow for reporting issues, prioritizing them based on their impact, and assigning them to responsible parties for resolution. Effective data quality issue management is essential as it enables Entities to swiftly address data inaccuracies, minimize operational disruptions, and maintain the integrity of their data. By implementing a robust issue management framework, Entities can ensure timely remediation of data quality problems, support regulatory compliance, and foster trust in their data assets. This process not only enhances data reliability but also promotes a proactive approach to data governance, driving overall effectiveness and success.

#### Data Quality Issue Management Key Benefits:

- **Promotes Proactive Issue Detection:** Implementing a structured process for data quality issue management enables early detection and swift resolution of data quality issues, thereby reducing the risk of data-related errors and enhancing overall data integrity.
- **Ensures Data Reliability:** Systematic identification and remediation of data quality issues ensure that the data remains accurate, consistent, and reliable, supporting better decision-making across the Entity.
- **Enhances Trust and Credibility:** Reliable data enhances the trust and credibility of the Entity among stakeholders, including customers, partners, and regulatory bodies, by demonstrating a commitment to high data quality standards.
- **Strengthens Compliance:** Effective data quality issue management ensures adherence to regulatory standards and internal policies, mitigating the risk of non-compliance and associated penalties.
- **Facilitates Continuous Improvement:** A robust issue management framework fosters a culture of continuous improvement by regularly identifying, analyzing, and resolving data quality issues, leading to ongoing enhancements in data processes.
- **Supports Informed Decision-Making:** High-quality data that is accurate and reliable underpins informed decision-making, leading to better strategic outcomes and enhanced organizational performance.

The remaining sections of this chapter will detail the below process that will guide the Entity on executing data quality issue management:

#### 5.1 Data Quality Issue Management and Remediation Process

## Guiding Principles for Data Quality Issue Management

1. **Establish Clear Issue Prioritization Criteria:** Develop criteria to prioritize data quality issues based on their impact on business processes, ensuring that the most critical issues are addressed first to minimize disruption.
2. **Conduct Thorough Root Cause Analysis:** Implement comprehensive root cause analysis procedures to identify and address the underlying causes of data quality issues, preventing recurrence and ensuring long-term data integrity.
3. **Ensure Accountability:** Define and assign clear roles and responsibilities for data quality issue management, fostering accountability and ensuring that all team members understand their specific duties in maintaining data quality.
4. **Maintain Transparency:** Promote open communication about data quality issues, their impacts, and remediation efforts. Transparency ensures that all stakeholders are informed and can collaborate effectively to resolve issues.
5. **Foster Continuous Improvement:** Encourage a culture of continuous improvement by regularly reviewing data quality management processes and incorporating feedback. Learn from past issues to enhance future data quality practices.
6. **Apply Consistent Standards:** Implement consistent data quality standards and procedures across the Entity. Consistency ensures that all departments adhere to the same criteria for managing and resolving data quality issues.
7. **Promote Collaboration:** Engage various departments in data quality issue management to leverage diverse expertise. Collaboration enhances problem-solving and ensures comprehensive solutions to data quality problems.
8. **Document Issues and Resolutions:** Maintain detailed records of data quality issues, analysis, and resolutions. Documentation supports accountability, transparency, and provides valuable insights for continuous improvement.
9. **Align with Compliance Requirements:** Ensure that data quality issue management practices comply with relevant regulatory requirements and internal policies, minimizing the risk of non-compliance and associated penalties.
10. **Integrate Quality into Workflows:** Embed data quality checks and processes into existing business workflows to make data quality a fundamental part of operations, ensuring ongoing attention to data integrity.

## 5.1 Data Quality Issue Management and Remediation Process

Effective management and remediation of data quality issues are critical for maintaining reliable and accurate data across the Entity. This section outlines the processes for identifying, prioritizing, and addressing data quality issues as they arise. By implementing a structured approach to issue management, Entities can quickly detect data anomalies, assess their impact, and take appropriate corrective actions. The goal is to minimize data defects, improve data consistency, and ensure that data remains a trusted asset for decision-making and operational processes. This process also emphasizes collaboration among data stewards, business units, and IT teams to ensure timely and effective remediation, fostering a proactive data quality culture.

### 5.1.1 Identify Data Quality Issues

This activity will guide the entity on the steps to identify data quality issues.

#### 5.1.1.A

**Identify Data Quality Issues:** The first step in data quality issue remediation is the identification process.

Data quality issues are identified through a combination of automated tools and manual processes, typically involving the following approaches:

- **Data Profiling:** Automated tools scan data to detect anomalies like missing values, duplicates, and incorrect formats, providing an initial assessment of data quality.
- **Rule-Based Validation:** Predefined data quality rules are applied to identify violations, such as missing mandatory fields or invalid data formats.
- **Manual Audits:** Data stewards or analysts manually review data samples to spot issues that automated tools may miss.
- **User Feedback:** End-users report discrepancies or errors they encounter while using the data, helping to identify issues not detected by automated processes.
- **External Data Comparison:** Data is cross-checked with external sources (e.g., address databases) to identify discrepancies and outdated information.

### RACI Matrix – Identify Data Quality Issues

*Indicative*

| Sr No.  | Activity                     | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Quality Specialist | Business Data Analyst | Data Engineer | Data Steward | Data Owner |
|---------|------------------------------|---------------------------------------|---------------------------------------------|----------------------------|-------------------------|-----------------------|---------------|--------------|------------|
| 5.1.1.A | Identify Data Quality Issues | I                                     | I                                           | A                          | R                       | R                     | I             | R            | I          |

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

## 5.1.2 Analyze and Prioritize Data Quality Issues

This activity will guide the entity on the steps to assess and prioritize data quality issues, ensuring that the most critical problems are addressed first to minimize disruption to business processes.

### 5.1.2.A

**Assess Impact:** Begin by evaluating the severity of each data quality issue, considering how it affects critical business processes and decision-making. Issues that significantly impact operations or regulatory compliance should be given higher priority. This assessment helps allocate resources effectively to the most pressing problems.

### 5.1.2.B

**Create a Prioritization List:** Develop a detailed list categorizing issues from high to low priority based on the impact assessment. This list should be regularly updated as new issues arise or as the impact of existing issues change. Prioritizing issues ensures that the Entity focuses on resolving the most critical problems first.

The prioritization criteria is as follows:

- **Impact:** How the issue affects business operations and decision-making
- **Urgency:** The timeframe needed for resolution
- **Frequency:** How often the issue occurs or affects data quality
- **Data Sensitivity:** The sensitivity of the data involved
- **Complexity:** The level of difficulty in resolving the issue (e.g., required resources, dependencies)

### Asset Guidance

The **Data Quality Issue Management Register Template** serves as a structured framework for identifying, documenting, and managing data quality issues within an Entity. It centralizes tracking of issues, prioritizes them based on severity and impact, and fosters accountability by assigning ownership.

*Illustrative*

| Issue ID | Issue Description                                   | Date Raised | Raised By         | Priority Level | Impact                                           | Status      | Assigned To       | Action Taken                                       | Resolution Date |
|----------|-----------------------------------------------------|-------------|-------------------|----------------|--------------------------------------------------|-------------|-------------------|----------------------------------------------------|-----------------|
| 11       | Duplicate customer records identified in CRM.       | 9/10/2023   | Data Quality Team | High           | Affects customer communication and reporting.    | Resolved    | Data Engineer     | Initiated deduplication process.                   | 9/13/2023       |
| 12       | Missing product prices in the inventory database.   | 2/3/2024    | Business Analyst  | Medium         | Delays in order processing and fulfillment.      | Resolved    | Inventory Manager | Contacted suppliers for missing info.              | 2/5/2024        |
| 13       | Customer email addresses not in valid format.       | 6/2/2024    | Customer Support  | High           | Affects communication and marketing outreach.    | In Progress | IT Support        | Created a validation rule in the system.           |                 |
| 14       | Outdated transaction dates in reporting system.     | 9/4/2024    | Business User     | Low            | Minor impact on reporting accuracy.              | In Progress | Data Analyst      | Updated records and adjusted reporting parameters. |                 |
| 15       | Inconsistent product descriptions across platforms. | 9/16/2024   | Marketing Team    | Medium         | Confusion among customers leading to complaints. | Open        | Product Manager   | Started cross-platform review.                     |                 |

| Notes/Comments                                     | Follow-Up Actions                                | Impact | Urgency | Frequency | Data Sensitivity | Complexity | Score | Priority Level |
|----------------------------------------------------|--------------------------------------------------|--------|---------|-----------|------------------|------------|-------|----------------|
| Need to assess impact on marketing campaigns.      | Implement regular deduplication checks.          | 1      | 2       | 1         | 3                | 2          | 1.80  | Medium         |
| Highlighted need for better data entry validation. | Establish stricter data entry rules.             | 2      | 3       | 2         | 1                | 3          | 2.05  | Medium         |
| Need to notify impacted customers.                 | Train staff on data entry standards.             | 2      | 1       | 1         | 1                | 1          | 1.35  | Low            |
| Reviewed reporting processes for accuracy.         | Schedule regular data audits.                    | 3      | 2       | 2         | 3                | 3          | 2.80  | High           |
| Need to ensure consistency for branding.           | Develop a centralized product description guide. | 2      | 2       | 1         | 3                | 3          | 2.35  | Medium         |

## RACI Matrix – Prioritize and Analyze Data Quality Issues

*Indicative*

| Sr No.  | Activity                     | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Strategy Expert | Data Quality Specialist | Business Data Analyst | Data Owner | Data Steward |
|---------|------------------------------|---------------------------------------|---------------------------------------------|----------------------------|----------------------|-------------------------|-----------------------|------------|--------------|
| 5.1.2.A | Assess Impact                | I                                     | I                                           | A                          | I                    | R                       | R                     | I          | R            |
| 5.1.2.B | Create a Prioritization List | I                                     | I                                           | A                          | C                    | R                       | R                     | I          | R            |

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

### 5.1.3 Formulate Solutions and Implement Remediation Actions

This activity will guide the entity on the steps to develop and execute effective remediation actions, addressing both technical and non-technical root causes to restore data quality.

#### Process – Formulate Solutions and Implement Remediation Actions

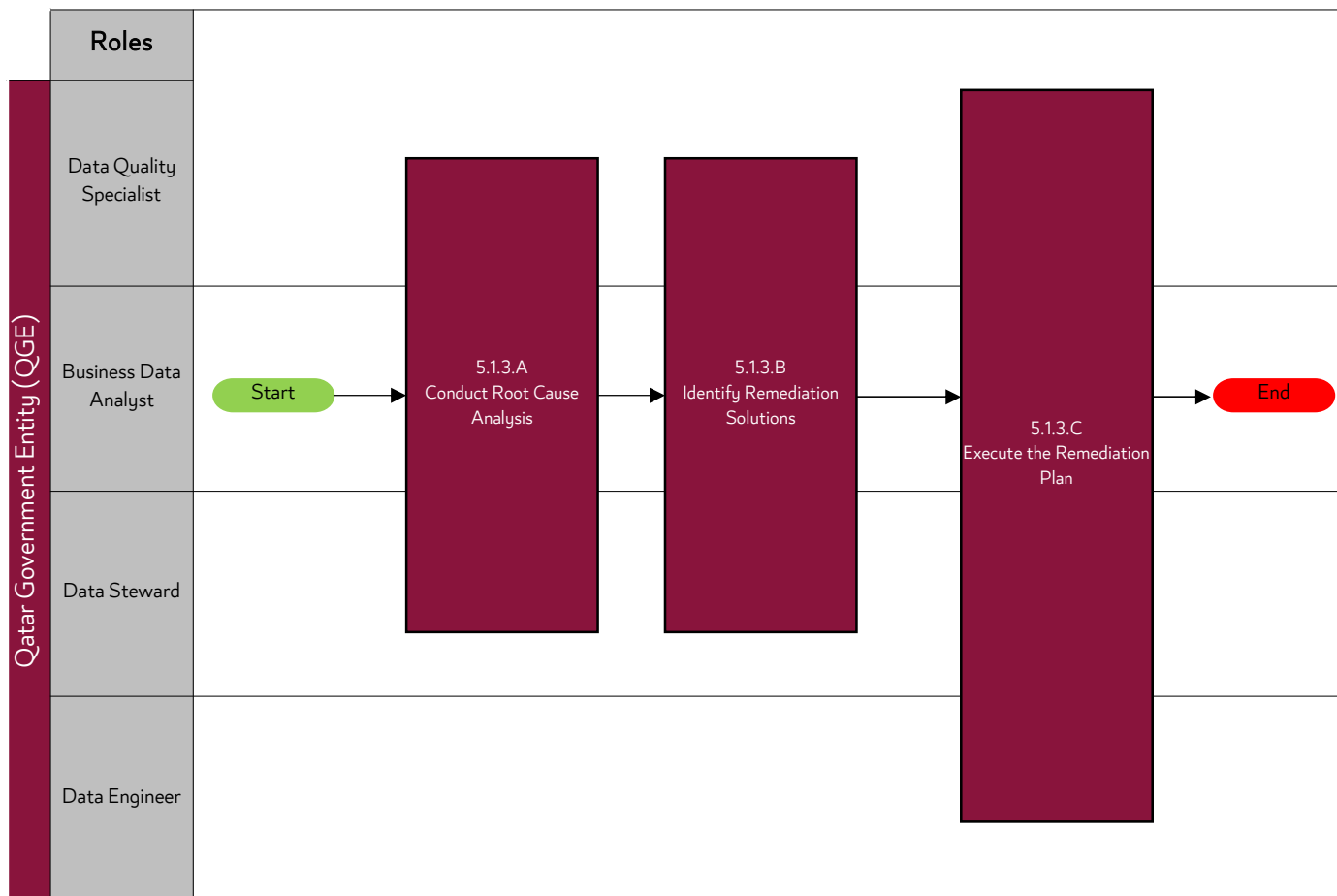


Figure 6: Formulate Solutions and Implement Remediation Actions Process Map

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

#### 5.1.3.A

**Conduct Root Cause Analysis:** Conduct a thorough root cause analysis to identify the underlying causes of data quality issues. This involves examining symptoms, tracing the data's journey through systems (data lineage), and distinguishing between technical (e.g., system errors) and non-technical causes (e.g., training gaps). Understanding the root causes is essential for developing effective remediation strategies. Additionally, this assessment helps allocate resources effectively to the most pressing problems.

#### 5.1.3.B

**Identify Remediation Solutions:** Create tailored remediation options that address both technical (e.g., updating systems, refining data validation rules, or automating data entry processes) and non-technical (e.g., enhanced training programs, increased leadership support, and improved accountability measures) root causes and develop a timeline with realistic deadlines for each solution.

## 5.1.3.C

***Execute the Remediation Plan:*** Implement the remediation actions according to the plan, closely monitoring progress to ensure adherence to the timeline. Engage all responsible parties and maintain open communication to address any obstacles promptly. Effective execution requires collaboration and commitment from all involved teams.

 **Asset Guidance**

A **Data Quality Remediation Plan** is a structured and systematic approach designed to identify, prioritize, and address data quality issues within an entity. This plan outlines specific actions and strategies along with implementation timelines aimed at correcting data inaccuracies, inconsistencies, and other quality deficiencies to ensure data integrity and reliability.

The process map below details the data quality issue analysis, remediation, and escalation process:

### Process – DQ Issue Remediation Process Map

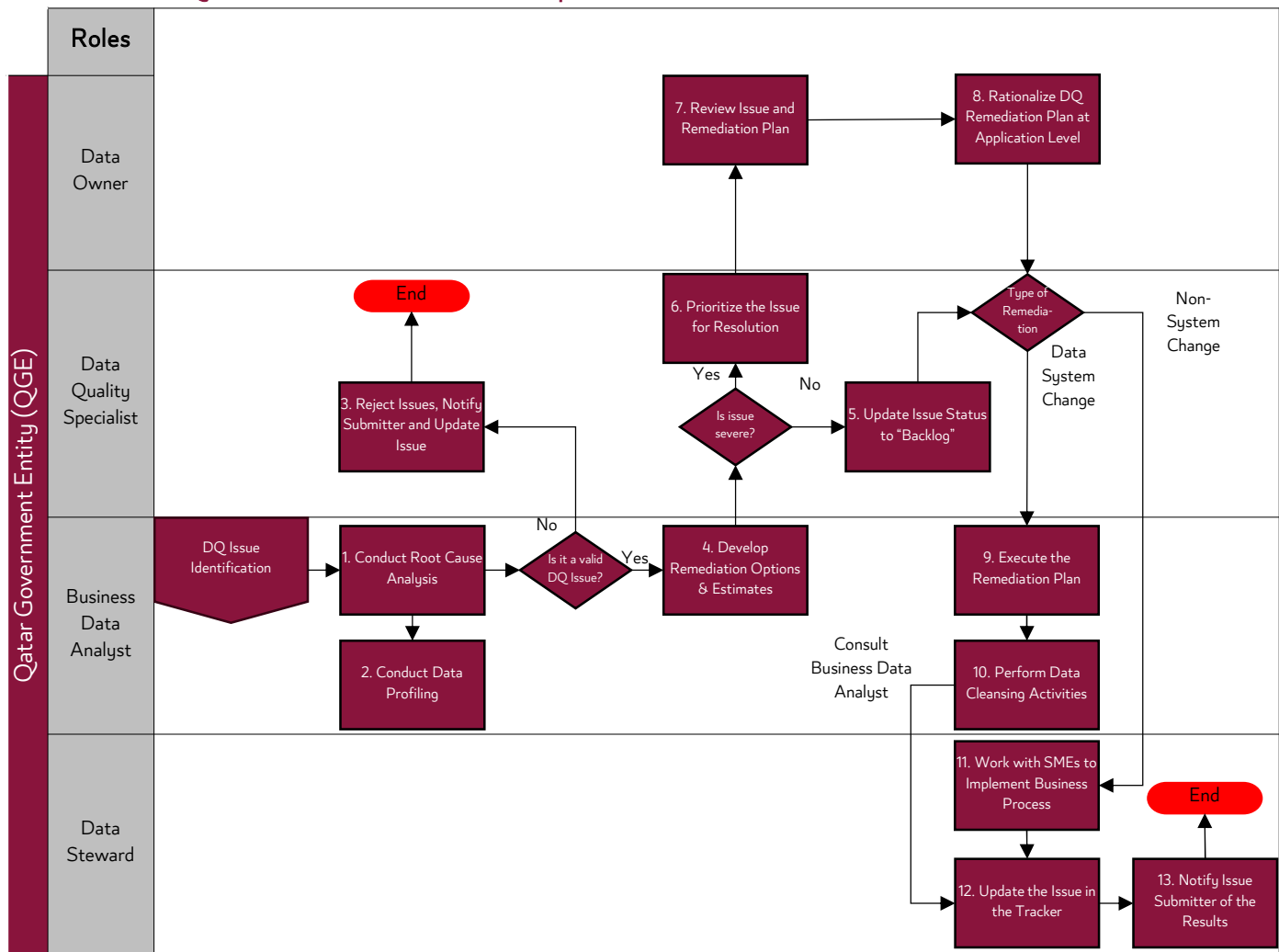


Figure 7: DQ Issue Remediation Process Map

### RACI Matrix – Formulate Solutions and Implement Remediation Actions

*Indicative*

| Sr No.  | Activity                       | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Strategy Expert | Data Quality Specialist | Business Data Analyst | Data Engineer | Data Owner | Data Steward |
|---------|--------------------------------|---------------------------------------|---------------------------------------------|----------------------------|----------------------|-------------------------|-----------------------|---------------|------------|--------------|
| 5.1.3.A | Conduct Root Cause Analysis    |                                       |                                             | A                          | C                    | R                       | R                     |               | I          | R            |
| 5.1.3.B | Identify Remediation Solutions | I                                     | I                                           | A                          | C                    | R                       | R                     | C             | I          | R            |
| 5.1.3.C | Execute the Remediation Plan   | I                                     | I                                           | A                          | I                    | R                       | R                     | R             | I          | R            |

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

### 5.1.4 Monitor and Evaluate Remediation Efforts

This activity will guide the entity on the steps to monitor and evaluate the progress and effectiveness of remediation efforts, ensuring that data quality issues are resolved satisfactorily.

5.1.4.A

**Track Progress:** Use data quality monitoring tools and systems to track the progress of remediation efforts in real time. Regular updates and progress reports can help identify any deviations from the plan early, allowing for quick adjustments. Monitoring ensures that the remediation actions stay on track and are completed as scheduled. In the absence of dedicated monitoring tools, scheduled manual data sampling or automated scripts can suffice and the results can be reported and visualized in Excel.

5.1.4.B

**Evaluate Effectiveness:** After implementing remediation actions, assess the actions' effectiveness in resolving the data quality issues. This involves verifying that the issues have been corrected, and that data quality has improved. Evaluating effectiveness helps ensure that the root causes have been adequately addressed and that similar issues will not recur.

5.1.4.C

**Fine-tune Remediation Plan:** If the initial remediation efforts do not fully resolve the issues, make necessary adjustments to the plan. This might involve revisiting the root cause analysis, refining solutions, or reallocating resources. Flexibility and willingness to adapt are crucial for successful data quality issue management.

#### RACI Matrix – Monitor and Evaluate Remediation Efforts

*Indicative*

| Sr No.  | Activity                   | Data Management and Security Director | Data Management and Governance Section Head | Data Quality Specialist | Data Owner | Data Steward |
|---------|----------------------------|---------------------------------------|---------------------------------------------|-------------------------|------------|--------------|
| 5.1.4.A | Track Progress             | I                                     | A                                           | R                       | I          | R            |
| 5.1.4.B | Evaluate Effectiveness     | I                                     | A                                           | R                       |            |              |
| 5.1.4.C | Fine-tune Remediation Plan | I                                     | A                                           | R                       |            | R            |

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

### 5.1.5 Document and Communicate

This activity will guide the entity on the steps to document the entire issue management and remediation process and maintain open communication with stakeholders, fostering transparency and collaboration.

5.1.5.A

**Record Details:** Maintain comprehensive documentation of the entire issue management and remediation process. This includes details of identified issues, root cause analysis findings, remediation actions taken, and outcomes. Thorough documentation supports transparency and provides a valuable reference for future data quality initiatives.

5.1.5.B

**Maintain Logs:** Keep detailed logs of all data quality issues and their resolution processes. These logs should be easily accessible for audit purposes and continuous improvement efforts. Documenting the process helps track the history of data quality issues and the effectiveness of remediation efforts.

5.1.5.C

**Stakeholder Communication:** Regularly update relevant stakeholders on the progress of issue management and remediation efforts. Provide clear and concise updates to ensure stakeholders are informed and can offer input if necessary. Effective communication fosters collaboration and ensures that all parties are aligned in their efforts to maintain data quality.

#### RACI Matrix – Document and Communicate

*Indicative*

| Sr No.  | Activity                  | Data Management and Security Director | Data Management and Governance Section Head | Data Quality Specialist | Data Owner | Data Steward |
|---------|---------------------------|---------------------------------------|---------------------------------------------|-------------------------|------------|--------------|
| 5.1.5.A | Record Details            |                                       | A                                           | R                       | I          | R            |
| 5.1.5.B | Maintain Logs             |                                       |                                             | A/R                     |            |              |
| 5.1.5.C | Stakeholder Communication | I                                     | A                                           | R                       |            | R            |

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

Based on this Activity, the Entity can



Prioritize and Analyze Data Quality Issues



Implement the Quality Remediation Plan



Monitor and Review Remediation Efforts

## 6. Data Quality Performance Management



## 6. Data Quality Performance Management

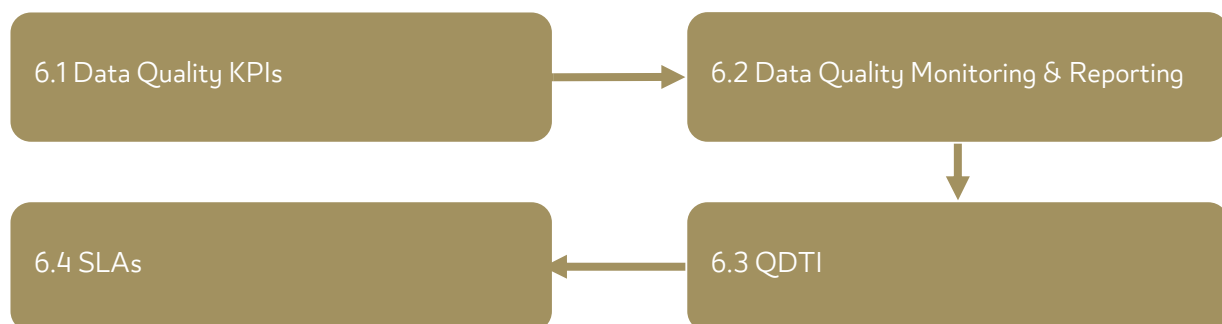
### What is Data Quality Performance Management and why does the Entity need it?

Data Quality Performance Management is a critical aspect of maintaining and improving the overall quality of an Entity's data. This involves establishing key performance indicators (KPIs) to monitor, measure, and report on data quality trends and the effectiveness of the data quality issues resolution process. By implementing service-level agreements (SLAs), the Entity can ensure that data quality objectives are consistently met and that performance standards are maintained. This systematic approach enables the Entity to track progress, identify areas for improvement, and drive continuous enhancement of data quality practices.

#### Data Quality Performance Management Key Benefits:

- **Improved Issue Resolution Efficiency:** Establishing and tracking performance metrics for the data quality issues resolution process ensures that problems are addressed promptly and effectively, reducing the time and resources spent on remediation.
- **Consistent Quality Standards:** Implementing SLAs helps maintain consistent data quality standards across the Entity, ensuring that all departments adhere to agreed-upon benchmarks and expectations.
- **Proactive Identification of Trends:** Regular measurement and reporting on data quality trends enable the Entity to proactively identify and address emerging issues before they escalate, minimizing potential disruptions.
- **Informed Strategic Decisions:** Performance management data provides valuable insights into the effectiveness of data quality initiatives, informing strategic decisions and helping prioritize future investments in data quality improvement.
- **Accountability and Transparency:** Clear KPIs and SLAs establish accountability for data quality across the Entity, fostering a culture of transparency and continuous improvement.

The remaining sections of this chapter will detail the below process that will guide the Entity on implementing data quality performance management:



## Guiding Principles for Data Quality Performance Management

1. **Align KPIs with Strategic Goals:** Ensure that key performance indicators (KPIs) are directly linked to the Entity's strategic objectives, supporting the overall mission. This alignment clarifies the purpose of data quality efforts and fosters collaboration across departments, helping everyone understand their contributions to the Entity's success.
2. **Focus on Critical Data Elements:** Select KPIs that target the most significant data elements affecting business operations. By concentrating resources on these critical areas, organizations can make informed decisions more efficiently, improving data quality and enhancing overall operational effectiveness.
3. **Establish Robust Monitoring and Reporting Mechanisms:** Implement real-time monitoring tools to quickly identify and resolve data quality issues, minimizing disruptions. Regular, structured reporting and automation enhance visibility, accountability, and accuracy in data quality management processes.
4. **Set and Maintain Service-Level Agreements (SLAs):** Create clear SLAs that define expectations for data quality performance and issue resolution timelines. Regularly track compliance and update SLAs to adapt to changing business needs and regulatory standards, ensuring they remain effective and relevant.
5. **Encourage Feedback and Collaboration:** Foster a culture of open communication across departments to gather feedback on data quality issues. Collaborative efforts drive continuous improvement and innovation, helping to identify challenges and enhance data quality initiatives.
6. **Invest in Training and Benchmarking:** Provide ongoing training and resources to ensure staff understand data quality principles and maintain high standards. Additionally, benchmark performance against industry best practices to learn from leaders and stay competitive.
7. **Assign Responsibilities and Conduct Regular Reviews:** Clearly define roles and responsibilities for data quality management to ensure accountability. Implement regular reviews and audits to ensure compliance with standards, identify improvement opportunities, and reinforce a culture of continuous enhancement.

## 6.1 Data Quality KPIs

Data quality Key Performance Indicators (KPIs) provide a measurable framework to assess and track the effectiveness of data quality management efforts. Examples of these KPIs include data quality levels across established dimensions such as accuracy, completeness, and consistency, as well as the number of data quality rules defined and their compliance rates. Additionally, KPIs can track the volume and percentage of data quality issues reported and resolved, as well as the average age of unresolved issues. Monitoring the number and percentage of escalated data quality issues is also crucial for identifying critical challenges that need immediate attention. These KPIs help organizations ensure continuous improvement in data quality, promote accountability, and support data-driven decision-making.

### 6.1.A

***Establish and Calculate Data Quality KPIs:*** Define specific data quality KPIs that enable executive stakeholders to evaluate the effectiveness of the data quality plan. These KPIs should provide insights into metrics such as accuracy, completeness, consistency, and timeliness, helping to assess how well the data quality initiatives align with and support overall business goals.

Calculate the performance measurement KPIs for the data quality program:

#### **Data Quality Standards:**

- # and % of data elements with meta-data (definitions).
- # of data quality business rules.
- # and % of data elements with metrics and thresholds.

#### **Data Quality Issue Resolution:**

- # of data quality issues reported within an Entity.
- # and % of data quality issues resolved.
- # and % of data quality issue rejected by Data Quality Team after issue analysis.
- Average age of open data quality issues (in days).
- # and % of escalated data quality issues.

#### **Data Quality Reports:**

- # of data quality scorecards/reports scheduled.
- # and % data domains covered under data quality reports.

#### **Data Quality Efforts:**

- # Resources (full-time and part-time).
- # Hours spent.



#### **Key Message / Information**

The Entity must establish a clear process for regularly reviewing and updating Data Quality KPIs to ensure their continued alignment with organizational goals and evolving data quality needs. This process should include setting defined intervals for KPI evaluation, incorporating stakeholder feedback, and assessing the effectiveness of each KPI in measuring key data quality metrics such as accuracy, completeness, and timeliness.

6.1.B

**Compare KPI Values:** Compare KPI values to historical trends to enable analysis of data quality performance. Build trend monitoring and detection into Data Quality performance reporting to be able to understand the improvements taking place over time and whether any corrective actions are needed.

### RACI Matrix – Data Quality KPIs

*Indicative*

| Sr No. | Activity                                  | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Quality Specialist | Data Owner | Data Steward |
|--------|-------------------------------------------|---------------------------------------|---------------------------------------------|----------------------------|-------------------------|------------|--------------|
| 6.1.A  | Establish and Calculate Data Quality KPIs | I                                     | A                                           | R                          | R                       | I          | C            |
| 6.1.B  | Compare KPI Values                        | I                                     | A                                           | R                          | R                       | I          | C            |

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

Based on this Activity, the Entity can



Establish KPIs for Data Quality Assessment



Monitor & Report on Data Quality KPIs

## 6.2 Data Quality Monitoring & Reporting

The Entity should monitor and report on the Entity's data quality based on the defined Data Quality rules and KPIs. This should include, at a minimum:

- Data quality dashboards to track, analyze, and display data in an aggregated format to gain deeper insight on the overall performance and trends across data quality KPIs
- Data quality scorecard report providing a high-level view comparison of data quality rules with the thresholds, highlighting areas which do not meet the thresholds (e.g.: with red-yellow-green colors)

Optionally, the reporting tool may also support the following advanced capabilities:

- Generating process control reports, primarily for use by data engineers. The reports will include custom reports (e.g.: for monitoring quality checks during data processing, based on operational metadata) or quality gate reports (e.g.: for capturing key business KPIs and ensuring data consistency across various data layers in the data warehouse or data lake)
- Generating automated Alerts and Notifications by implementing a system/ dashboard that automatically alerts data stewards and owners when data quality issues are detected or when certain thresholds are breached

### Process – Data Quality Monitoring & Reporting

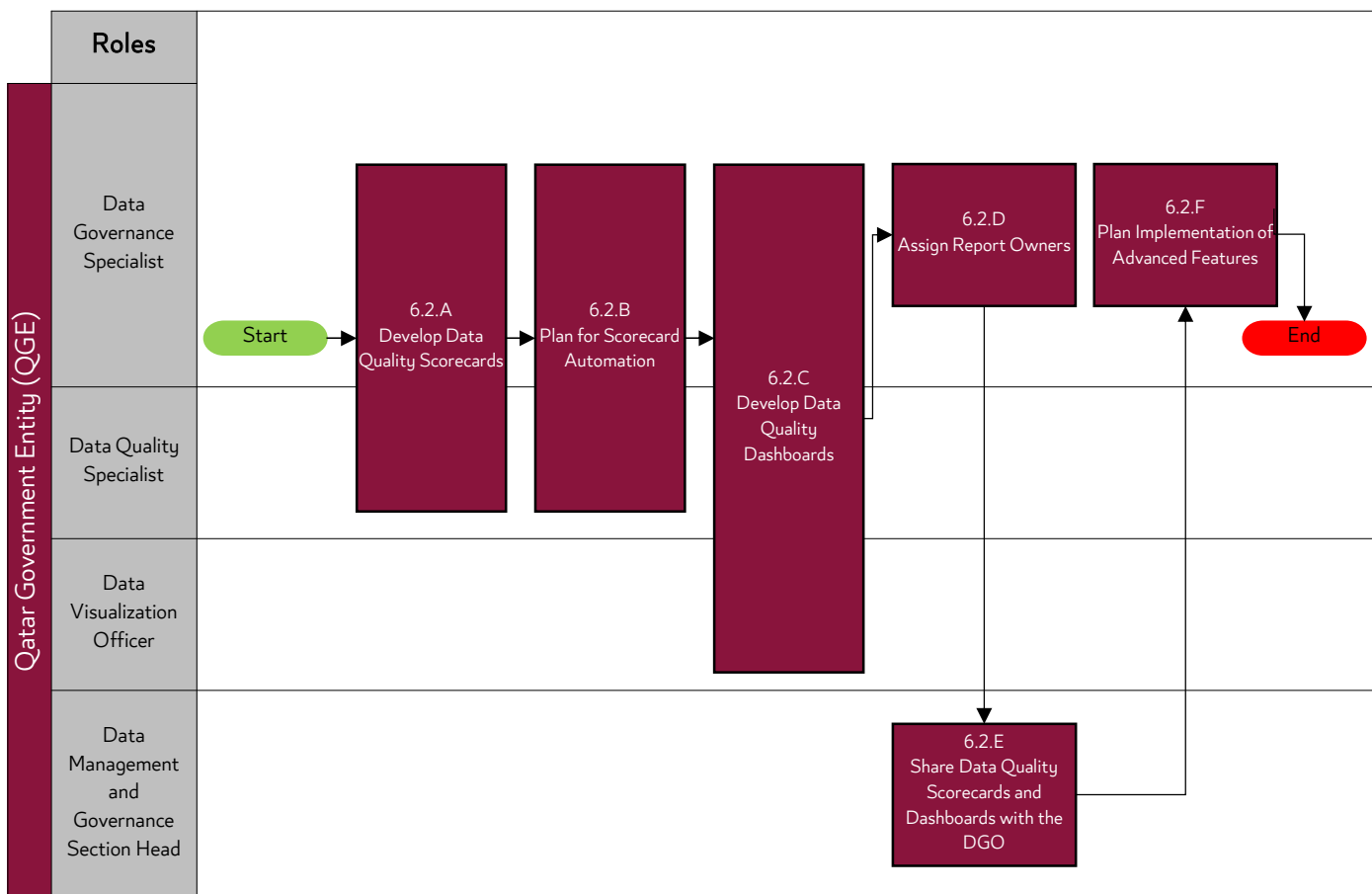


Figure 8: Data Quality Monitoring & Reporting Process Map

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

## 6.2.A

**Develop Data Quality Scorecards:** Scorecards are used for periodic monitoring of data quality status of the CDEs within the different data domains and sub domains. The scorecard will have the following characteristics:

- Each CDE has a separate score for each data quality dimension.
- Each CDE has an overall data quality score based on its aggregated Data Quality score across the Data Quality dimensions.
- Data domain and subdomain overall score can be then calculated based on aggregating the Data Quality scores for the CDEs considering the weight that is assigned based on its critically to business from 1 to 10, where 10 is critical and 1 is low.
- Data owner assigns the CDE weight.

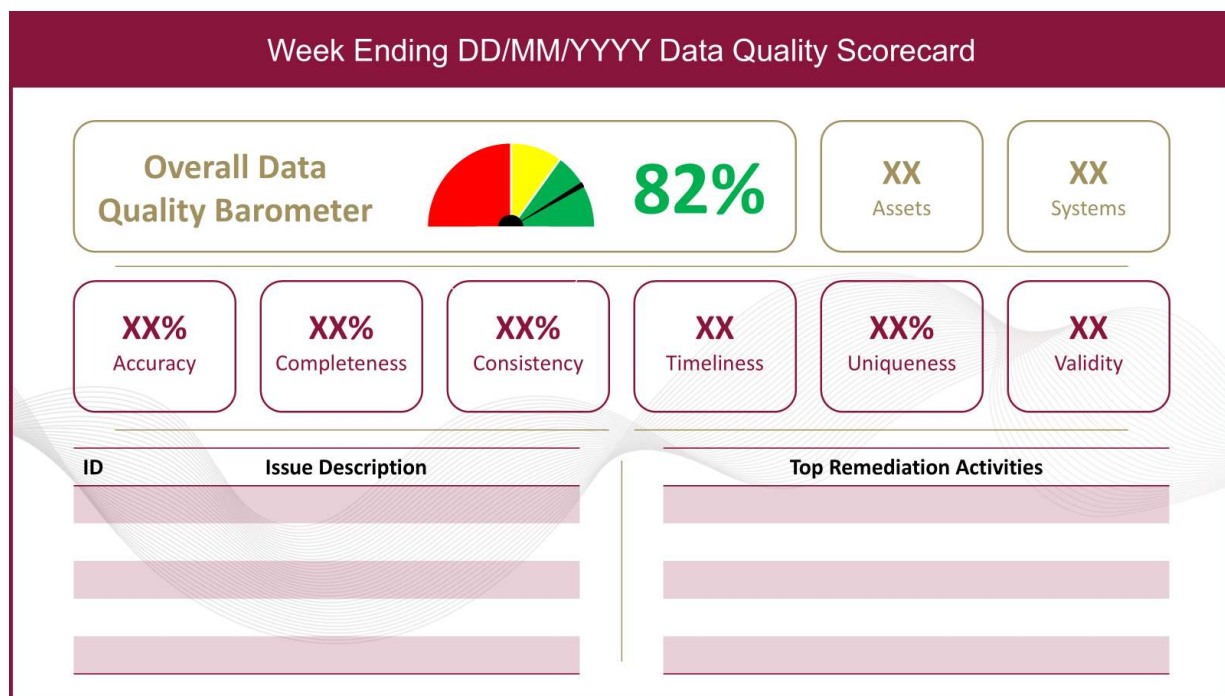
## 6.2.B

**Plan for Scorecard Automation:** Once the scorecards are developed and the population process is clear, consider automating the flows to reduce overhead and improve reporting efficiency. This can be done through generating and running scheduled scripts and outputting the result onto a visualization tool.

### 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*



## 6.2.C

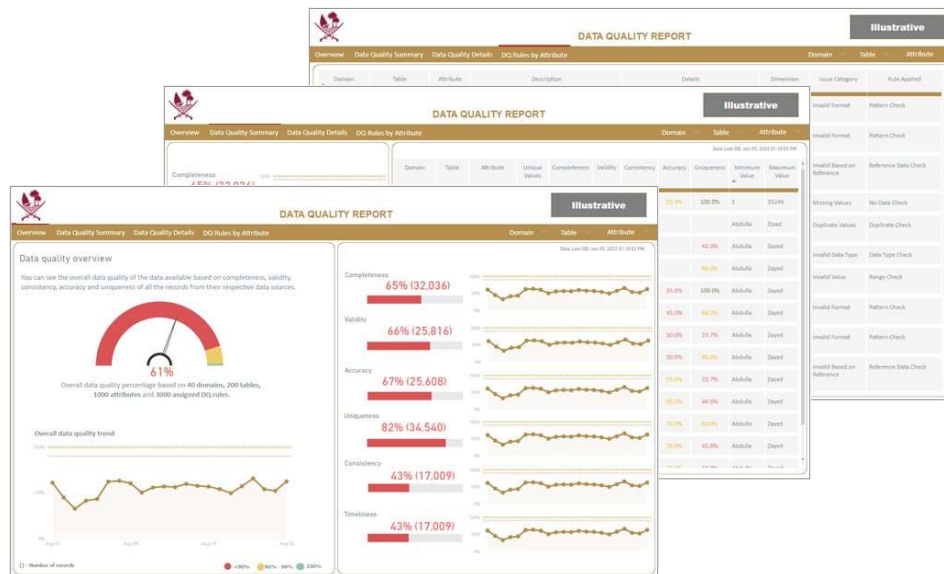
**Develop Data Quality Dashboards:** Develop an Executive data quality Dashboard with the following characteristics:

- **Executive Summary View:** The dashboard should provide a high-level overview of data quality metrics, enabling executives to quickly assess overall performance without delving into operational details.
- **Aggregated Insights:** It should aggregate data from multiple sources, presenting summarized scores for key data quality dimensions such as completeness, accuracy, and validity, ensuring a holistic view of data health.
- **Visual Clarity:** Utilize clear visualizations, such as bar charts and pie charts, to convey trends and insights intuitively, making it easy for executives to interpret complex data quickly.
- **Key Performance Indicators (KPIs):** Display essential KPIs that reflect the effectiveness of data quality initiatives, allowing executives to monitor performance and make informed strategic decisions.
- **Trend Analysis:** Include historical comparisons and trend visualizations to help executives understand how data quality has evolved over time, facilitating better forecasting and resource allocation.

### 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*



6.2.D

**Assign Report Owners:** Assign an owner to each report or dashboard to check the results periodically and communicate any raised issues.

6.2.E

**Share Data Quality Scorecards and Dashboards with the DGO:** Share the Data Quality Scorecards and Dashboards with the National Data Governance Office (DGO) monthly, enabling the DGO to monitor progress, track improvements, and ensure alignment with National Data Standards.

6.2.F

**Plan Implementation of Advanced Features:** After successfully implementing the monitoring and reporting measures with the fundamental and essential functionalities, set a timeline and plan for implementation of advanced, nice-to-have features in the future to support the long-term strategic objectives of data quality management.

### RACI Matrix – Data Quality Reporting & Monitoring

*Indicative*

| Sr No. | Activity                                                  | DGO | Data Management and Security Director | Data Management and Governance Section Head | Data Governance Specialist | Data Quality Specialist | Business Data Analyst | Data Visualization Officer | Data Owner | Data Steward |
|--------|-----------------------------------------------------------|-----|---------------------------------------|---------------------------------------------|----------------------------|-------------------------|-----------------------|----------------------------|------------|--------------|
| 6.2.A  | Develop Data Quality Scorecards                           |     | I                                     | A                                           | R                          | R                       | C                     | C                          | I          | C            |
| 6.2.B  | Plan for Scorecard Automation                             |     | I                                     | A                                           | R                          | R                       |                       |                            |            |              |
| 6.2.C  | Develop Data Quality Dashboards                           |     | I                                     | A                                           | R                          | R                       | C                     | R                          | I          | C            |
| 6.2.D  | Assign Report Owners                                      |     | I                                     | A                                           | R                          |                         |                       |                            |            |              |
| 6.2.E  | Share Data Quality Scorecards and Dashboards with the DGO | I   | A                                     | R                                           |                            |                         |                       |                            |            |              |
| 6.2.F  | Plan Implementation of Advanced Features                  |     | I                                     | A                                           | R                          |                         |                       |                            |            |              |

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

Based on this Activity, the Entity can



Build Data Quality Dashboards and Reports

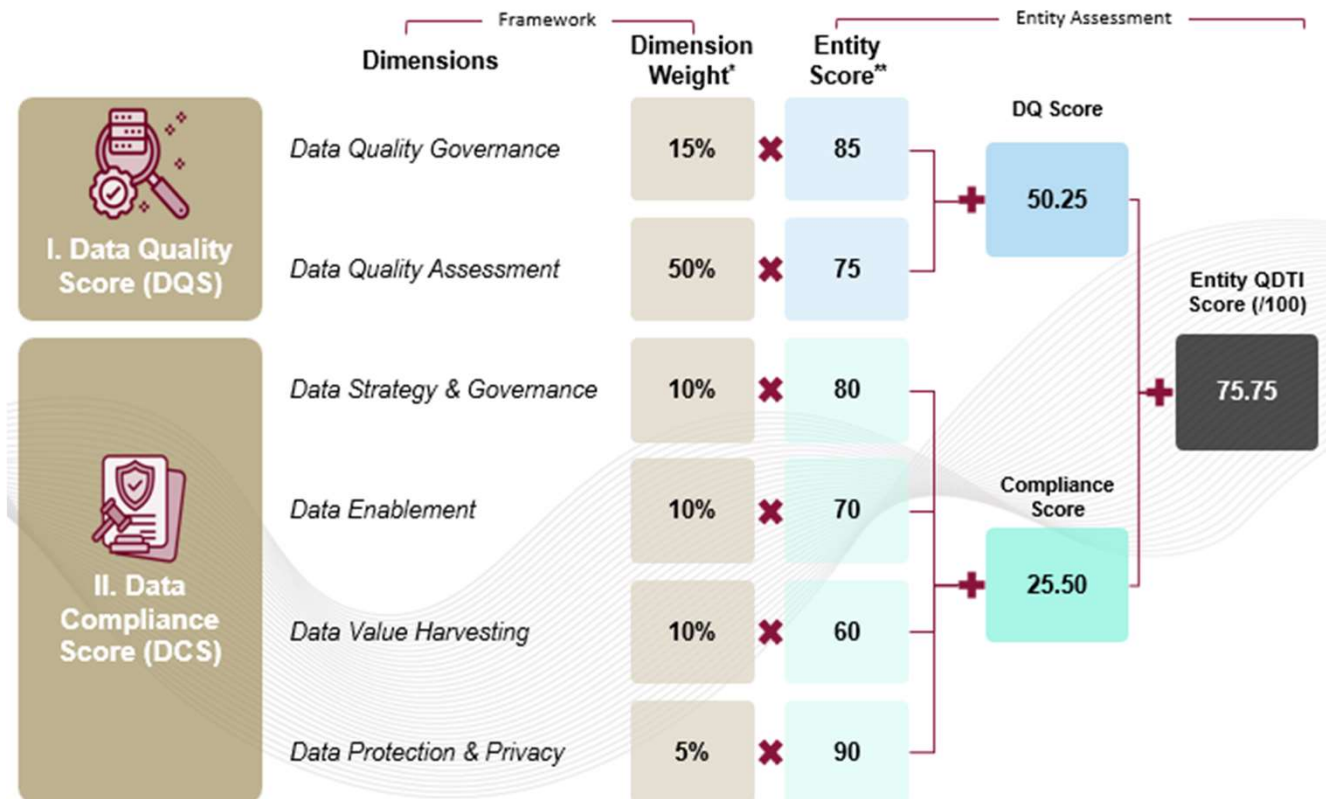


Monitor Data Quality Levels and Performance

## 6.3 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 DQS.



**\*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 9: Illustrative QDTI Scoring Methodology for an Entity

### 6.3.A

**Initiate Data Quality Score Calculation:** 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 Data Quality Management Standards as well as reporting the Data Quality Assessment results.

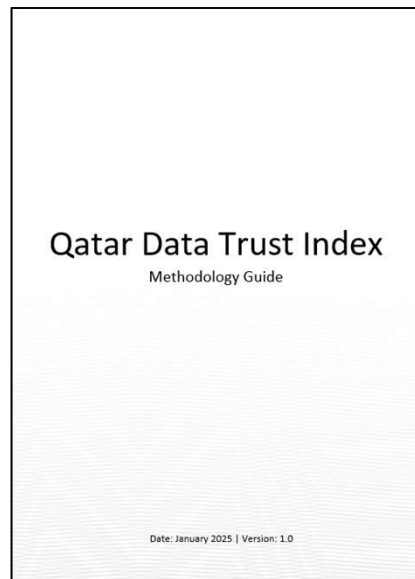
## 6.3.B

**Conduct Data Quality Governance Assessment:** Initiate a self-assessment to evaluate adherence to the National Data Quality Management 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 and record the Data Quality Assessment results from previous exercises. 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 NDQAF 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.

## 6.3.C

**Pull Data Quality Assessment Results and Report to DGO:** Retrieve the results from the previously conducted data quality assessment exercise, including performance metrics and supporting evidence. These results provide valuable insights into the Entity's historical and ongoing data quality performance. Submit the retrieved results, along with the necessary evidence, to the DGO to contribute to the calculation of the overall Data Quality Score (DQS).

## 6.3.D

**DGO to Calculate the Data Quality Score (DQS):** The DGO will analyze the Data Quality Governance Assessment results and the Data Quality Assessment results to calculate the Data Quality Score (DQS) for the Entity. The DQS will reflect the Entity's adherence to Data Quality Management Standards and its overall data quality performance.



## 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 | Data Quality Specialist |
|--------|--------------------------------------------------------|-----|-------------------------------------|---------------------------------------------|----------------------------|----------------------------|-------------------------|
| 6.3.A  | Initiate Data Quality Score Calculation                |     | I                                   | A                                           | R                          |                            |                         |
| 6.3.B  | Conduct Data Quality Governance Assessment             | I   | I                                   | A                                           | R                          | C                          | C                       |
| 6.3.C  | Pull Data Quality Assessment Results and Report to DGO |     | I                                   | A                                           | R                          | C                          | C                       |
| 6.3.D  | DGO to Calculate the Data Quality Score (DQS)          | A/R | I                                   | I                                           | I                          | I                          | I                       |

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

Based on this Activity, the Entity can



Calculate Data Quality Score (DQS)

## 6.4 Data Quality SLAs

The Entity should establish and implement service-level agreements (SLAs) for business-critical data quality issues. The SLAs should include, at a minimum:

- Data elements covered by the agreement
- Business impacts associated with data flaws
- Data Quality dimensions and business rules associated with each data element
- Methods for measuring against those expectations
- Acceptability threshold for each measurement
- Steward(s) to be notified in case the acceptability threshold is not met
- Timelines and deadlines for expected resolution or remediation of the issue
- Escalation process when the SLA is not met

Optionally, the Entity may enhance its data quality SLAs capabilities as follows:

- Automate the monitoring of SLAs by implementing a system that automatically monitor compliance with SLAs
- Generate automated notifications and actions by implementing systems that automatically trigger emails or notifications to relevant stakeholders if SLAs are breached

### Process – Data Quality SLAs

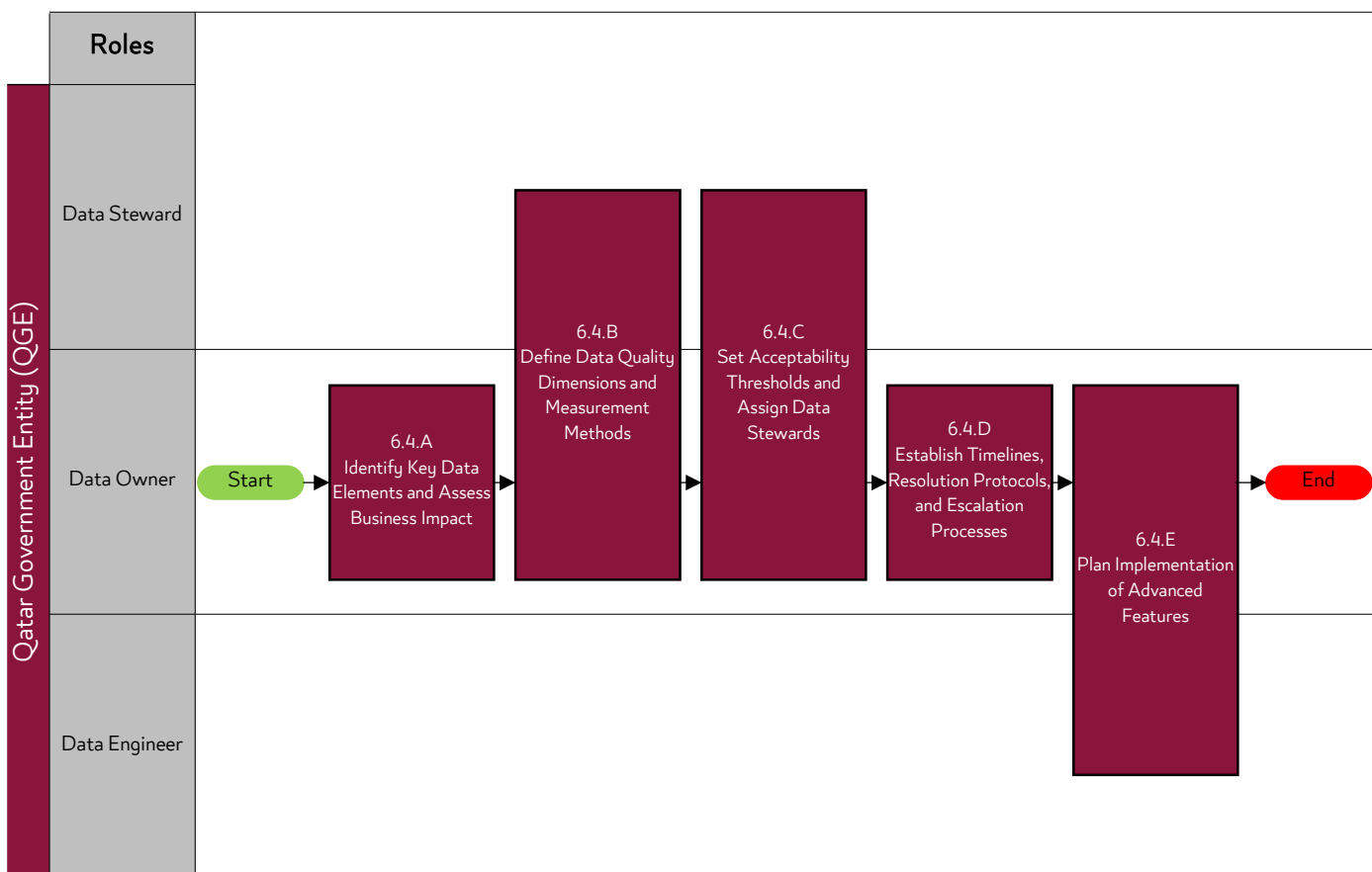


Figure 10: Data Quality SLAs Process Map

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

## 6.4.A

**Identify Key Data Elements and Assess Business Impact:** Determine which business-critical data elements will be covered by the SLAs and analyze the business impacts associated with data flaws for each. Engage stakeholders to ensure alignment on what constitutes critical data.

## 6.4.B

**Define Data Quality Dimensions and Measurement Methods:** In alignment with the National Data Quality Framework; establish specific data quality dimensions (e.g., Completeness, Consistency, Uniqueness) along with associated business rules. Outline methods for measuring compliance against these expectations, specifying tools and metrics for regular assessment.

## 6.4.C

**Set Acceptability Thresholds and Assign Data Stewards:** Define acceptable performance thresholds for each measurement, ensuring they align with business goals. Designate data stewards responsible for monitoring compliance and addressing any issues when thresholds are not met.

## 6.4.D

**Establish Timelines, Resolution Protocols, and Escalation Processes:** Set clear timelines for the expected resolution of data quality issues and define escalation processes when SLAs are not met. This ensures timely responses and accountability.

## 6.4.E

**Plan Implementation of Advanced Features:** Set a timeline and plan for implementation of advanced, nice-to-have features in the future to support the long-term strategic objectives of data quality management. Consider automating the monitoring of SLAs to track compliance in real-time and implement notification systems for breaches. Regularly review and update the SLAs to keep them relevant and effective, incorporating stakeholder feedback.

## RACI Matrix – Data Quality SLAs

*Indicative*

| Sr No. | Activity                                                            | Data Strategy Expert | Data Quality Specialist | Data Engineer | Data Architect | Data Owner | Data Steward |
|--------|---------------------------------------------------------------------|----------------------|-------------------------|---------------|----------------|------------|--------------|
| 6.4.A  | Identify Key Data Elements and Assess Business Impact               | C                    | A                       |               |                | R          | C            |
| 6.4.B  | Define Data Quality Dimensions and Measurement Methods              | C                    | A                       |               |                | R          | R            |
| 6.4.C  | Set Acceptability Thresholds and Assign Data Stewards               | C                    | A                       |               |                | R          | R            |
| 6.4.D  | Establish Timelines, Resolution Protocols, and Escalation Processes | C                    | A                       |               |                | R          | C            |
| 6.4.E  | Plan Implementation of Advanced Features                            |                      | A                       | R             | C              | R          | C            |

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

Based on this Activity, the Entity can



Develop Data Quality SLAs

## 7. Appendix



## 7.1 Templates & Frameworks

| Asset No. | Asset Name                                                                | Description                                                                                                                                                  | File Name                                                                |
|-----------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| 1         | Data Quality Policy Template                                              | Template containing a sample table of contents to refer to when drafting the Entity's policy.                                                                | NPC_DGO_QDKC_DQ_Data Quality Policy Template                             |
| 2         | Data Management Policy Communication Email Template                       | Template containing a sample email used to announce the publishing of the Entity's new policy.                                                               | NPC_DGO_QDKC_General_Data Management Policy Communication Email Template |
| 3         | Data Profiling Report Template                                            | Template used for summarizing data profiling results, including data quality assessment, statistical summaries, and insights on data patterns and anomalies. | NPC_DGO_QDKC_DQ_Data Profiling Report Template                           |
| 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 Quality Critical Data Elements Prioritization Framework and Template | Template to help Entities prioritize CDEs for effective data quality management.                                                                             | NPC_DGO_QDKC_DQ_CDE Prioritization Framework and Template                |
| 6         | 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                |
| 7         | Data Quality Rules Template                                               | Template used to establish data quality rules and thresholds within an Entity to enforce data quality standards.                                             | NPC_DGO_QDKC_DQ_Data Quality Rules Template                              |
| 8         | Data Quality Tool Selection Criteria Template                             | Template containing key factors for evaluating and selecting data quality tools, covering functionality, integration, cost and other factors.                | NPC_DGO_QDKC_DQ_Data Quality Tool Selection Criteria Template            |
| 9         | Data Quality Assessment Template                                          | Template used to provide entities with a structured approach to evaluating the quality of their datasets across multiple dimensions.                         | NPC_DGO_QDKC_DQ_Data Quality Assessment Template                         |
| 10        | Issue Management Register Template                                        | Template used for logging and tracking data quality issues, including fields for issue identification, description, priority, status, and resolution steps.  | NPC_DGO_QDKC_DQ_Issue Management Register Template                       |

| Asset No. | Asset Name                               | Description                                                                                                                                               | File Name                                     |
|-----------|------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|
| 11        | Data Quality Remediation Plan            | Framework outlining steps to correct data quality issues, including root cause analysis, remediation strategies, resource allocation, and timelines.      | NPC_DGO_QDKC_DQ_Data Quality Remediation Plan |
| 12        | 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        |
| 13        | Qatar Data Trust Index Methodology Guide | Guide that can be leveraged by the Entity when conducting the compliance self-assessment.                                                                 | 22012025_NPC_QDTI Methodology Guide_V1.0      |
| 14        | QDTI Assessment Template                 | Template that can be used to document compliance with the National Data Standards and record the Data Quality Assessment results from previous exercises. | 23012025_NPC_QDTI Assessment Template_V1.0    |



المجلس الوطني للتخطيط  
National Planning Council

دولة قطر • State of Qatar