

Data Quality Management Plan

Source File: generated-documents\dmbok\data-quality-management-plan.md

Generated: 30/07/2025 at 06:57:43

Generated by: Requirements Gathering Agent - PDF Converter

Data Quality Management Plan

Generated by adpa-enterprise-framework-automation v3.2.0

Category: dmbok

Generated: 2025-07-16T09:30:42.461Z

Description: Defines the organization's approach to data quality management, including objectives, standards, roles, responsibilities, and processes in alignment with DMBOK best practices.

Data Quality Management Plan

Project:

Name: adpa-enterprise-framework-automation

Description: Modular, standards-compliant Node.js/TypeScript automation framework for enterprise requirements, project, and data management. Provides CLI and API for BABOK v3, PMBOK 7th Edition, and DMBOK 2.0 (in progress). Production-ready Express.js API with TypeSpec architecture. Designed for secure, scalable, and maintainable enterprise automation.

Version: 3.2.0

1. Introduction

1.1 Purpose

This Data Quality Management Plan (DQMP) establishes the framework, processes, and responsibilities for ensuring high data quality throughout the lifecycle of the adpa-enterprise-framework-automation project. The plan’s objective is to maintain the integrity, reliability, and usability of all data managed or produced by the system, supporting regulatory compliance (BABOK, PMBOK, DMBOK), effective automation, and business decision-making.

1.2 Scope

This plan applies to all data assets managed by the framework, including project requirements, templates, user information, documents, metadata, integration artifacts (e.g., Adobe, SharePoint), and audit logs, across all environments (development, staging, production).

1.3 Objectives

- Define data quality standards and metrics
 - Assign clear roles and responsibilities for data quality
 - Establish processes for assessing, monitoring, and improving data quality
 - Enable transparency through reporting and issue tracking
-

2. Data Quality Roles and Responsibilities

Role	Responsibilities
Data Owners	Accountable for data assets within their domain; approve changes and resolve escalated issues.

Role	Responsibilities
Data Stewards	Coordinate data quality activities; enforce standards; monitor and report on data quality metrics.
Data Quality Analysts	Profile, measure, and assess data; identify anomalies; recommend remediation actions.
Developers	Implement data validations in code and integration pipelines; remediate technical issues.
System Administrators	Support data backup, integrity, and security; restore/recover data as needed.
Business Users	Identify and report data quality issues; provide business context for data requirements.

Roles are mapped to specific team members in the RACI matrix (appendix), and responsibilities are integrated into the DevOps lifecycle.

3. Data Quality Dimensions

Dimension	Definition	Example (in context)
Accuracy	Data correctly describes the real-world entity or event.	Project requirement fields match actual stakeholder needs.

Dimension	Definition	Example (in context)
Completeness	All required data is present.	All mandatory fields in templates and API payloads are populated.
Consistency	Data is the same across various systems and formats.	Requirement status is consistent between API, CLI, and DB.
Timeliness	Data is up-to-date and available when needed.	Project status updates are reflected in API and user dashboards swiftly.
Uniqueness	Each data entity is recorded only once; no duplicates exist.	No duplicate user accounts or requirement IDs.
Validity	Data conforms to defined formats, ranges, and business rules.	Email addresses, GUIDs, and date fields adhere to format standards.

4. Data Quality Assessment

4.1 Data Profiling

- **Initial Profiling:** Conducted on onboarding new data sources, schemas, or integrations (e.g., Adobe API, SharePoint, external APIs).
- **Automated Profiling:** Leverage tools (e.g., express-validator, joi, zod, ajv) for schema validation and anomaly detection in API payloads.

4.2 Data Quality Monitoring

- **Continuous Monitoring:** Automated scripts and middleware (express-winston, morgan, winston) log data anomalies, validation errors, and exceptions.
- **Scheduled Reviews:** Data stewards review key data sets (templates, requirements, user lists) monthly for anomalies and quality trends.

4.3 Data Quality Measurement

- **Validation Rules:** Enforced at API, CLI, and DB layers using validation libraries (joi, zod, ajv) and OpenAPI/TypeSpec specifications.
 - **Test Automation:** Use Jest and custom test scripts (test:unit, test:performance) for regression and edge-case coverage.
 - **Data Sampling:** Periodic random sampling and manual review of records for qualitative assessment.
-

5. Data Quality Issue Management

5.1 Identification

- **Automated Detection:** Validation middleware and test suites flag data quality issues in real time.
- **User Reporting:** Issues reported via API error messages, user interface, or ticketing system.

5.2 Logging

- **Centralized Log Management:** All data quality incidents logged via winston and stored with metadata (timestamp, type, severity, origin).
- **Issue Tracker:** Use project issue tracker (e.g., GitHub Issues, Azure DevOps) for tracking open/closed status.

5.3 Remediation

- **Triage:** Data Quality Analysts and Stewards triage issues by severity and impact.
- **Resolution:** Remediation may involve data correction, code/config fixes, or process updates.
- **Root Cause Analysis:** Required for recurring or critical quality incidents.
- **Verification:** Post-remediation validation to confirm resolution before closure.

5.4 Escalation and Communication

- **Escalation Path:** Issues not resolved within SLAs are escalated to Data Owners and Project Managers.
- **Communication:** Regular updates provided to stakeholders and affected users.

6. Data Quality Tools and Technology

Tool/Library	Functionality
express-validator, joi, zod, ajv	API and payload validation, enforcing data schemas and business rules
winston, express-winston, morgan	Centralized logging for data quality events and anomalies
Jest, ts-jest	Automated testing, regression, and edge-case validation
Swagger, OpenAPI, TypeScript	API documentation and contract enforcement
Azure API Center/SharePoint	Data governance, integration validation, and metadata enforcement

Tool/Library	Functionality
Custom Scripts	Data profiling, batch validation, and migration checks

7. Data Quality Metrics and Reporting

7.1 Key Performance Indicators (KPIs)

- **Data Accuracy Rate:** % of records passing all validation checks
- **Completeness Score:** % of mandatory fields populated across datasets
- **Consistency Index:** % of records with aligned values across systems
- **Timeliness Metric:** Average latency between data update and availability
- **Uniqueness Ratio:** Number of duplicate records per dataset
- **Validity Rate:** % of fields conforming to format and business rules

7.2 Dashboards and Reporting

- **Automated Dashboard:** Web or CLI dashboards display real-time and historical KPIs, with drill-down to data anomalies.
- **Periodic Reports:** Monthly summary reports distributed to Data Owners and Project Managers.
- **Alerts:** Automated notifications (e.g., via email or Slack) for major data quality breaches or trends.

7.3 Continuous Improvement

- **Root Cause Analysis:** Incorporated into retrospectives for recurring issues.
 - **Lessons Learned:** Used to update validation rules, documentation, and training.
-

Appendix:

- RACI Matrix (roles-to-person assignment)
 - Data asset inventory and criticality assessment
 - Reference to DMBOK 2.0, BABOK v3, PMBOK 7th Edition standards
-

Approval:

This plan is reviewed and approved by the Data Governance Board and is subject to annual review or upon significant architectural changes.

End of Data Quality Management Plan
