

SAP Business Data Cloud and Enterprise AI Data Quality Assurance Framework

A Comprehensive Framework for Trusted, Governed, Semantically Consistent, AI-Ready Data Quality Across SAP Business Data Cloud and Enterprise AI

Sanjeeve Kumar Gajadi*

*(HCL Technologies, Bengaluru, Karnataka, India)

Email: sanjeevekumar.g@hcltech.com

Abstract:

Enterprise AI depends on trusted data, but organizations often manage data quality separately from AI assurance and business semantics. SAP Business Data Cloud creates opportunities to unify governed SAP and non-SAP data products, semantic context, analytics, and AI consumption. This paper proposes a SAP Business Data Cloud and Enterprise AI Data Quality Assurance Framework that connects data quality dimensions, semantic governance, lineage, ownership, data contracts, observability, AI readiness, and business outcomes. The framework covers completeness, accuracy, consistency, timeliness, validity, uniqueness, integrity, semantic alignment, and access quality. It defines quality rules, scorecards, certification, issue management, root-cause analysis, AI-use-case thresholds, and continuous monitoring across source, transformation, data-product, semantic, and consumption layers. The proposed approach enables enterprises to improve trust in SAP Business Data Cloud data products and reduce AI errors caused by poor or misunderstood enterprise data.

Keywords — SAP Business Data Cloud, Data Quality, Enterprise AI, SAP Datasphere, Data Products, Data Governance, Semantic Layer, Data Lineage, AI Readiness, Data Observability

I. INTRODUCTION

AI quality cannot exceed the reliability and semantic correctness of the data used for retrieval, training, inference, and decision support.

SAP Business Data Cloud provides an enterprise context in which data products, semantics, governance, and AI consumption can be managed together.

II. FRAMEWORK OBJECTIVES

The framework aims to measure, certify, monitor, and improve data quality according to business and AI use-case requirements.

Quality should be owned jointly by business data owners, data-product teams, platform teams, and AI consumers.

III. DATA QUALITY DIMENSIONS

Core dimensions include completeness, accuracy, consistency, timeliness, validity, uniqueness, integrity, and semantic correctness.

AI use cases may also require freshness, representativeness, provenance, sensitivity labeling, and retrieval suitability.

IV. DATA PRODUCT QUALITY

Each data product should define quality expectations, ownership, service levels, critical fields, validation rules, and consumer obligations.

Quality contracts provide a stable interface between producers and AI or analytics consumers.

V. SEMANTIC QUALITY

Semantic quality ensures that business definitions, measures, hierarchies, and relationships are

interpreted consistently across applications and AI systems.

Misaligned semantics can create plausible but incorrect AI outputs even when source data values are technically accurate.

VI. LINEAGE AND ROOT CAUSE

End-to-end lineage should connect source systems, transformations, data products, semantic models, and AI consumption.

Lineage enables faster root-cause analysis when an AI outcome is traced to stale, missing, or incorrectly transformed data.

VII. QUALITY RULES AND SCORECARDS

Rules can be executed at ingestion, transformation, publication, and consumption stages.

Scorecards should show current quality, trends, breached thresholds, affected consumers, and accountable owners.

VIII. AI READINESS CERTIFICATION

Data products intended for AI should satisfy minimum thresholds for quality, lineage, sensitivity, representativeness, and access control.

Certification should be risk-based and renewed when material source or transformation changes occur.

IX. DATA OBSERVABILITY

Observability should detect schema change, volume anomalies, freshness failures, failed transformations, unexpected nulls, and semantic drift.

Alerts should connect technical failures to affected business capabilities and AI use cases.

X. GOVERNANCE AND ISSUE MANAGEMENT

Quality issues require severity, business impact, owner, remediation target, root cause, and closure evidence.

Recurring issues should drive improvement of source processes and data contracts rather than repeated downstream fixes.

XI. INTEGRATION WITH AI LIFECYCLE

AI evaluations should reference certified data versions and record the data products used during testing and production.

Material data-quality deterioration should trigger model or agent reassessment when it can affect outputs.

XII. CONCLUSION

The SAP Business Data Cloud and Enterprise AI Data Quality Assurance Framework connects trusted data products with AI lifecycle assurance.

By combining quality rules, semantics, lineage, certification, observability, and ownership, enterprises can improve the reliability of AI decisions built on SAP and non-SAP data.

ACKNOWLEDGMENT

The author acknowledges the contributions of enterprise architecture practitioners, AI governance specialists, standards bodies, security professionals, data leaders, and technology researchers whose published work informed the concepts synthesized in this paper.

REFERENCES

- [1] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0), 2023.
- [2] National Institute of Standards and Technology, AI RMF Generative Artificial Intelligence Profile, 2024.
- [3] ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system.
- [4] ISO 31000:2018, Risk management — Guidelines.
- [5] ISO/IEC 27001:2022, Information security, cybersecurity and privacy protection — Information security management systems.
- [6] The Open Group, TOGAF Standard, 10th Edition, 2022.
- [7] ISACA, COBIT 2019 Framework: Governance and Management Objectives, 2019.

- [8] OWASP Foundation, OWASP Top 10 for Large Language Model Applications.
- [9] MITRE, ATLAS: Adversarial Threat Landscape for Artificial-Intelligence Systems.
- [10] OECD, Recommendation of the Council on Artificial Intelligence.
- [11] SAP SE, SAP Business Technology Platform Documentation.
- [12] SAP SE, SAP Business Data Cloud and SAP Datasphere Documentation.