

# Enterprise Data Product Architecture Framework

A Comprehensive Enterprise Architecture for Reusable, Governed, Discoverable, Interoperable, and AI-Ready Data Products

Sanjeeve Kumar Gajadi\*

\*(HCL Technologies, Bengaluru, Karnataka, India)

Email: [sanjeevekumar.g@hcltech.com](mailto:sanjeevekumar.g@hcltech.com)

\*\*\*\*\*

## Abstract:

Modern enterprises are shifting from project-centric data pipelines toward reusable data products that are owned, governed, discoverable, and consumed through standardized interfaces. This transformation requires an architecture that connects business domains, semantic models, data contracts, quality, lineage, metadata, security, interoperability, observability, lifecycle governance, and measurable value. This paper proposes an Enterprise Data Product Architecture Framework that defines the structural and operating principles required to design and manage data products across SAP and non-SAP landscapes. The framework integrates SAP Business Data Cloud, SAP Datasphere, SAP HANA Cloud, APIs, event-driven integration, metadata catalogs, data quality services, data sharing, analytics, and AI consumption. It establishes product roles, architecture layers, data contracts, governance controls, service levels, lifecycle stages, and KPIs/KRIs. The proposed approach enables organizations to reduce duplication, improve trust and reuse, accelerate analytics and Artificial Intelligence, and strengthen business accountability for enterprise data.

**Keywords** — Data Product, Enterprise Data Architecture, SAP Business Data Cloud, SAP Datasphere, Data Mesh, Data Governance, Data Contracts, Data Quality, Lineage, AI-Ready Data

\*\*\*\*\*

## I. INTRODUCTION

Traditional data programs often organize delivery around projects, reports, or pipelines. This creates duplicated datasets, inconsistent semantics, fragile dependencies, and unclear ownership.

A data-product approach treats trusted data as a reusable business capability with explicit consumers, contracts, service levels, governance, and lifecycle management.

## II. DATA PRODUCT PRINCIPLES

A data product should be discoverable, understandable, interoperable, trustworthy, secure, reusable, measurable, and owned by a business-aligned domain.

Products should expose stable interfaces and semantic meaning so consumers are insulated from unnecessary internal implementation changes.

## III. DOMAIN AND OWNERSHIP MODEL

Business domains should identify critical entities, events, measures, and analytical concepts for which they are accountable.

Each product requires a business owner, data steward, technical owner, and defined consumer responsibilities.

## IV. ARCHITECTURE LAYERS

The architecture includes source systems, ingestion and integration, storage and processing, semantic modeling, metadata and catalog, quality, security, product APIs, event services, analytics, and AI consumption.

Layer separation allows teams to evolve implementation technologies without breaking product contracts.

## V. DATA CONTRACTS

Data contracts should define schema, semantics, units, quality expectations, freshness, service levels, ownership, versioning, and compatibility rules.

Breaking changes should be managed through explicit versioning and consumer impact analysis.

## **VI. SEMANTIC AND BUSINESS MODEL**

Measures, dimensions, hierarchies, master data, and business definitions should be governed as reusable semantic assets.

Consistent semantics improve interoperability across analytics, planning, AI, and operational processes.

## **VII. QUALITY AND OBSERVABILITY**

Products should implement automated quality checks for completeness, validity, accuracy, consistency, timeliness, and uniqueness.

Observability should track freshness, failed pipelines, schema drift, usage, performance, and consumer incidents so teams can manage products against service commitments.

## **VIII. METADATA, CATALOG, AND LINEAGE**

A catalog should connect business descriptions, owners, schemas, quality indicators, classifications, policies, and lineage.

Lineage helps consumers understand provenance and enables teams to assess downstream impact when sources, transformations, or semantics change.

## **IX. SECURITY AND ACCESS**

Products should enforce least privilege, classification-based controls, row or column restrictions, masking, encryption, and purpose limitation.

Access decisions should apply consistently across APIs, SQL, analytics tools, data sharing, and AI retrieval.

## **X. SAP BUSINESS DATA CLOUD AND DATASPHERE**

SAP Business Data Cloud and SAP Datasphere can provide governed business data, semantic modeling, data sharing, and integration across SAP landscapes.

Spaces, connections, data flows, models, and business layers should follow enterprise standards for ownership, naming, transport, security, and lifecycle.

## **XI. INTEROPERABILITY AND INTEGRATION**

Data products should be consumable through SQL, APIs, events, files, or governed sharing depending on use case and latency needs.

SAP Integration Suite and event-driven services can support controlled exchange between operational systems and data products.

## **XII. AI-READY DATA PRODUCTS**

AI-ready products require stronger provenance, quality, context, freshness, permissions, and suitability metadata because AI can amplify hidden data defects.

Products used in RAG or agentic AI should preserve source authorization and expose machine-readable semantic context where possible.

## **XIII. PRODUCT LIFECYCLE**

The lifecycle includes ideation, design, build, certification, publication, operation, versioning, recertification, deprecation, and retirement.

Retirement should consider downstream dependencies and provide migration paths for active consumers.

## **XIV. OPERATING MODEL**

A federated operating model balances domain ownership with enterprise standards. Central teams can provide platform services, catalog, policy, quality tooling, and architecture guidance.

Domains remain accountable for product meaning, quality, and value.

## **XV. VALUE MEASUREMENT**

KPIs can include product reuse, active consumers, quality pass rate, delivery lead time, service-level compliance, and reduction in duplicate datasets.

KRIs can include unowned products, broken contracts, stale data, excessive access, critical quality failures, and undocumented dependencies.

## **XVI. IMPLEMENTATION ROADMAP**

Organizations should begin with a small number of high-value domains and products, establish common contracts and governance, and validate reuse.

The model can then scale through reusable templates, automated quality, metadata harvesting, policy-as-code, and portfolio management.

## **XVII. CONCLUSION**

Enterprise data products provide a practical architecture for turning data into a reusable, governed business capability.

The proposed framework connects ownership, contracts, semantics, quality, lineage, security,

interoperability, lifecycle, and AI readiness to support scalable analytics and digital transformation.

## REFERENCES

- [1] ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system.
- [2] ISO/IEC 27001:2022, Information security, cybersecurity and privacy protection — Information security management systems.
- [3] ISO 31000:2018, Risk management — Guidelines.
- [4] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0), 2023.
- [5] National Institute of Standards and Technology, AI RMF Generative Artificial Intelligence Profile, 2024.
- [6] National Institute of Standards and Technology, SP 800-207: Zero Trust Architecture, 2020.
- [7] The Open Group, TOGAF Standard, 10th Edition, 2022.
- [8] ISACA, COBIT 2019 Framework: Governance and Management Objectives, 2019.
- [9] SAP SE, SAP Business Technology Platform Documentation.
- [10] SAP SE, SAP AI Core and SAP AI Launchpad Documentation.
- [11] SAP SE, SAP Integration Suite Documentation.
- [12] SAP SE, SAP HANA Cloud Documentation.
- [13] SAP SE, SAP Cloud Identity Services Documentation.
- [14] SAP SE, SAP Cloud ALM Documentation.
- [15] SAP SE, SAP Business Data Cloud and SAP Datasphere Documentation.
- [16] OWASP Foundation, OWASP Top 10 for Large Language Model Applications.
- [17] MITRE, ATLAS: Adversarial Threat Landscape for Artificial-Intelligence Systems.
- [18] OECD, Recommendation of the Council on Artificial Intelligence.
- [19] IEEE, Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems.
- [20] DAMA International, DAMA-DMBOK2: Data Management Body of Knowledge, 2017.

## AUTHOR BIOGRAPHY

Sanjeeve Kumar Gajadi is an SAP Enterprise Architect and Lead Consultant with extensive experience in enterprise data architecture, SAP

Datasphere, SAP Business Data Cloud, SAP HANA, analytics, integration, governance, and AI-enabled transformation. His professional interests include data products, semantic architecture, data mesh, AI-ready data, and enterprise governance.

## ACKNOWLEDGMENT

The author acknowledges the contributions of enterprise architecture practitioners, international standards bodies, AI governance communities, security researchers, and SAP documentation teams whose published guidance informed the concepts synthesized in this paper.