Canonical Requirements

Generated: 2026-02-04T15:30:26

This PDF is generated from canonical requirements JSON files. Statements are rendered as the primary source-of-truth summary for downstream analysis.

1. Executive Summary

High-level purpose and outcomes; include only explicit scope/purpose facts from source.

System-System End Point Lineage Calculation

Source type: confluence-xhtml
Source: confluence/DataModel-195297281-System-System End Point Lineage Calculation.xhtml
Confluence pageId: 195297281
Confluence space: DataModel
Confluence version: 33
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/195297281/System-System+End+Point+Lineage+Calculation
Retrieved at: 2026-02-03T18:25:25.521006

REQ0000001 [Fact]

This document provides a high-level approach for calculating "lineage complete" based on a minimal amount of lineage map information and describes an algorithm for traversing the tree backwards and forwards.

Story	—
BDD	—

Practical Data: Data Governance at Scale - Data Lineage Capability & Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-199491586-Practical Data- Data Governance at Scale - Data Lineage Capability & Metadata Store.xhtml
Confluence pageId: 199491586
Confluence space: DataModel
Confluence version: 42
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/199491586/Practical+Data+Data+Governance+at+Scale+-+Data+Lineage+Capability+Metadata+Store
Retrieved at: 2026-02-03T18:25:24.045921

REQ0000001 [Fact]

This document is part of the Data Governance at Scale series and specifically focuses on the Data Lineage Capability and Metadata Store components.

Story	—
BDD	—

Practical Data: Data Governance at Scale - Conventions

Source type: confluence-xhtml
Source: confluence/DataModel-202735632-Practical Data- Data Governance at Scale - Conventions.xhtml
Confluence pageId: 202735632
Confluence space: DataModel
Confluence version: 2

REQ0000001 [Fact]

This document defines conventions used across the document series.

Story	—
BDD	—

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

REQ0000001 [Fact]

This document is part of a series proposing a target reference architecture for Data Governance at scale for large organizations with fragmented business processes and data flows, and provides a deeper dive into an architectural component.

Story	—
BDD	—

REQ0000002 [Fact]

Within each document section, the document covers metadata data model design and functional capability design, and provides links to other documents in the series.

Story	—
BDD	—

2. Scope and Context

Business context, drivers, and boundaries explicitly stated in sources.

(No items)

3. Assumptions and Constraints

All explicit assumptions and hard constraints.

System-System End Point Lineage Calculation

Source type: confluence-xhtml
Source: confluence/DataModel-195297281-System-System End Point Lineage Calculation.xhtml
Confluence pageId: 195297281
Confluence space: DataModel
Confluence version: 33
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/195297281/System-System+End+Point+Lineage+Calculation
Retrieved at: 2026-02-03T18:25:25.521006

REQ0000002 [Constraint]

The document does not establish a definition of "Authorised Source"; that definition is specific to organizational data governance policy.

Story	—
BDD	—

REQ0000003 [Assumption]

All definitions and calculations refer to attributes and tech-based lineage and do not account for context-based lineage.

Story	—
BDD	—

Practical Data: Data Governance at Scale - Data Lineage Capability & Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-199491586-Practical Data- Data Governance at Scale - Data Lineage Capability & Metadata Store.xhtml
Confluence pageId: 199491586
Confluence space: DataModel
Confluence version: 42
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/199491586/Practical+Data+Data+Governance+at+Scale+-+Data+Lineage+Capability+Metadata+Store
Retrieved at: 2026-02-03T18:25:24.045921

REQ0000004 [Assumption]

Data Inventory information (physical and business attributes) is stored within the metadata repository component.

Story	—
BDD	—

REQ0000005 [Assumption]

Data lineage information describes the flow of data within systems and across system boundaries; inventory discovery may identify additional dependencies to add to inventory.

Story	—
BDD	—

REQ0000006 [Assumption]

Technical lineage is defined as the relationship between two lowest-level attributes in the data inventory; higher-level dependencies can be implied.

Story	—
BDD	—

REQ0000007 [Assumption]

It may not always be possible to automatically discover technical lineage; a method is required for system applications to contribute lineage information manually.

Story	—
BDD	—

REQ0000008 [Constraint]

Data model designs are illustrative and graph-styled for convenience and do not require a specific storage method, database technology, or implementation design.

Story	—
BDD	—

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-202833967-Practical Data- Data Governance At Scale - Data Inventory Capability and Metadata Store.xhtml
Confluence pageId: 202833967
Confluence space: DataModel
Confluence version: 35
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202833967/Practical+Data+Data+Governance+At+Scale+-+Data+Inventory+Capability+and+Metadata+Store
Retrieved at: 2026-02-03T18:25:22.745224

REQ0000008 [Constraint]

Data inventorying requires access to data structure information (for example, fields and schemas).

Story	—
BDD	—

REQ0000009 [Constraint]

Data profiling requires access to underlying data content.

Story	—
BDD	—

REQ0000010 [Constraint]

The data model designs are for illustration and convenience only; although styled as a graph database with nodes and relationships, they do not require a specific data storage method, database technology, or implementation design.

Story	—
BDD	—

REQ0000012 [Constraint]

Implementations may use any number of levels and alternative names for technical inventory hierarchy levels; the only mandatory level is attribute, where all data lineage relationships are defined.

Story	—
BDD	—

REQ0000019 [Constraint]

Implementations may use any number of levels and alternative names for business inventory hierarchy levels.

Story	—
BDD	—

REQ0000024 [Constraint]

The business-logical model connects to the technical model at the lowest levels: most business attributes relate to a single technical attribute; a single business attribute may relate to multiple technical attributes; the reverse (one technical attribute relating to multiple business attributes) is not valid.

Story	—
BDD	—

REQ0000040 [Assumption]

Given the scale of inventory information, the majority of inventory information is expected to be discovered and updated automatically.

Story	—
BDD	—

4. Definitions and Glossary

All explicit definitions and term meanings.

System-System End Point Lineage Calculation

REQ0000004 [Definition]

End Point Value (EPV) is an attribute in the lineage that represents the end of the combined physical lineage as the desired business outcome.

Story	—
BDD	—

REQ0000005 [Definition]

Authoritative Source (AS) is an attribute representing the start of a physical lineage path that has been confirmed as an official source, requiring no further decomposition.

Story	—
BDD	—

REQ0000006 [Definition]

Each node must have a Status classification such as EPV, AS, Complete (COM), or Unknown (XX).

Story	—
BDD	—

REQ0000007 [Definition]

Backward depth from EPV sequences completeness roll-up; EPVs have backdepth=0 and backward depth is the minimum depth of downstream nodes plus 1, calculated per endpoint.

Story	—
BDD	—

REQ0000008 [Definition]

Forward depth from AS is the maximum hops from authoritative sources plus 1; authoritative sources have fordepth=0.

Story	—
BDD	—

REQ0000009 [Definition]

Node Score represents the current completeness of a node based on the average completeness of upstream nodes; AS nodes score 100%, nodes without lineage score 0%, and partial branches score between 0% and 100%.

Story	—
BDD	—

Practical Data: Data Governance at Scale - Conventions

Source type: confluence-xhtml
Source: confluence/DataModel-202735632-Practical Data- Data Governance at Scale - Conventions.xhtml
Confluence pageId: 202735632
Confluence space: DataModel
Confluence version: 2
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202735632/Practical+Data+Data+Governance+at+Scale+-+Conventions
Retrieved at: 2026-02-03T18:25:29.764240

REQ0000004 [Definition]

For NodeType INVENTORY and NodeSubType ATTRIBUTE, the naming convention is ATT and an example ID is ATT000000001.

Story	—
BDD	—

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-202833967-Practical Data- Data Governance At Scale - Data Inventory Capability and Metadata Store.xhtml
Confluence pageId: 202833967
Confluence space: DataModel
Confluence version: 35
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202833967/Practical+Data+Data+Governance+At+Scale+-+Data+Inventory+Capability+and+Metadata+Store
Retrieved at: 2026-02-03T18:25:22.745224

REQ0000011 [Definition]

Attribute is the lowest level of a technical data management hierarchy and represents the level at which data values are technically named, stored, and managed.

Story	—
BDD	—

REQ0000013 [Definition] — Application definition

Application (also known as a system) is the highest level of decomposition of a technical inventory of data elements.

Story	—
BDD	—

REQ0000014 [Definition] — Component definition

Component describes a sub-component of an application (for example, a database).

Story	—
BDD	—

REQ0000015 [Definition] — Schema definition

Schema is a logical subdivision of a component, often relating to a group of logical tables inside a database.

Story	—
BDD	—

REQ0000016 [Definition] — Table definition

Table is a logical grouping of related attributes associated with an individual record.

Story	—
BDD	—

REQ0000017 [Definition] — Attribute definition (hierarchy item)

Attribute is the lowest level of a technical hierarchy and represents the point at which data values are stored and managed.

Story	—
BDD	—

REQ0000018 [Definition]

Element is the lowest level of the business data management hierarchy and represents the level at which the business understands how the data is named and used.

Story	—
BDD	—

REQ0000020 [Definition] — Domain definition

Domain is the highest level at which business data is defined and owned.

Story	—
BDD	—

REQ0000021 [Definition] — Class definition

Class (or object) is the highest level at which the business understands a complete data object.

Story	—

BDD	—

REQ0000022 [Definition] — (Sub)Class definition

(Sub)Class is the next level of breakdown of a complete business object and represents a logical grouping of business elements; organizations may allow multiple subclass depths.

Story	—
BDD	—

REQ0000023 [Definition] — Element definition (hierarchy item)

Element is the lowest level of the business hierarchy and represents the point at which the business understands the data value.

Story	—
BDD	—

REQ0000025 [Definition]

For INVENTORY/ATTRIBUTE nodes in the Data Governance Metadata Repository, the nodeID field is the unique ID of the node within the repository.

Story	—
BDD	—

REQ0000026 [Definition]

For INVENTORY/ATTRIBUTE nodes, the nodeType field value is INVENTORY.

Story	—
BDD	—

REQ0000027 [Definition]

For INVENTORY/ATTRIBUTE nodes, the nodeSubType field value is ATTRIBUTE.

Story	—
BDD	—

REQ0000028 [Definition]

For INVENTORY/ATTRIBUTE nodes, nodeName is the technical (unique) name of the attribute within the technical hierarchy.

Story	—
BDD	—

REQ0000029 [Definition]

For INVENTORY/ATTRIBUTE nodes, nodeFQName is the derived fully-qualified unique name of the node within the technical hierarchy, derived from top to bottom (for example, ApplicationName/ComponentName/.../SchemaName/TableName/AttributeName).

Story	—

BDD	—

REQ0000031 [Definition]

For INVENTORY/ATTRIBUTE nodes, nodeSpecification is additional technical inventory information collected about a node as defined by the operating model.

Story	—
BDD	—

REQ0000032 [Definition]

For INVENTORY/ELEMENT nodes in the Data Governance Metadata Repository, the nodeID field is the unique ID of the node within the repository.

Story	—
BDD	—

REQ0000033 [Definition]

For INVENTORY/ELEMENT nodes, the nodeType field value is INVENTORY.

Story	—
BDD	—

REQ0000034 [Definition]

For INVENTORY/ELEMENT nodes, the nodeSubType field value is ELEMENT.

Story	—
BDD	—

REQ0000035 [Definition]

For INVENTORY/ELEMENT nodes, nodeName is the individual business name of the data element.

Story	—
BDD	—

REQ0000036 [Definition]

For INVENTORY/ELEMENT nodes, nodeFQName is the derived fully-qualified unique name of the node within the technical hierarchy, derived from top to bottom.

Story	—
BDD	—

5. Business Requirements

Business-level requirements and rules.

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-202833967-Practical Data- Data Governance At Scale - Data Inventory Capability and Metadata Store.xhtml
Confluence pageId: 202833967
Confluence space: DataModel
Confluence version: 35
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202833967/Practical+Data+Data+Governance+At+Scale+-+Data+Inventory+Capability+and+Metadata+Store
Retrieved at: 2026-02-03T18:25:22.745224

REQ0000004 [Fact]

Data Inventory information is the most fundamental information about data storage and flows within the organization and serves as the basis for other aspects of data governance, including lineage and quality.

Story	—
BDD	—

6. Functional Requirements

System behaviors and functional capabilities.

System-System End Point Lineage Calculation

Source type: confluence-xhtml
Source: confluence/DataModel-195297281-System-System End Point Lineage Calculation.xhtml
Confluence pageId: 195297281
Confluence space: DataModel
Confluence version: 33
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/195297281/System-System+End+Point+Lineage+Calculation
Retrieved at: 2026-02-03T18:25:25.521006

REQ0000010 [Requirement]

The algorithm must calculate node depth by working backwards from EPVs and then calculate node score by working forwards toward EPVs.

Story	—
BDD	—

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-202833967-Practical Data- Data Governance At Scale - Data Inventory Capability and Metadata Store.xhtml
Confluence pageId: 202833967
Confluence space: DataModel
Confluence version: 35
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202833967/Practical+Data+Data+Governance+At+Scale+-+Data+Inventory+Capability+and+Metadata+Store
Retrieved at: 2026-02-03T18:25:22.745224

REQ0000003 [Requirement]

The Data Inventory Capability must enable the collection and management of technical and business inventory data elements across the organization.

Story	—
BDD	—

REQ0000005 [Requirement]

Data Inventory capabilities must be able to scan and inventory systemic and business data elements and store them in the Data Governance Metadata store.

Story	—
BDD	—

REQ0000006 [Fact]

At a minimum, Data Inventory discovery enables the collection of data elements.

Story	—
BDD	—

REQ0000007 [Fact]

Additional Data Profiling capabilities can enable automated classification and description of data based on its content, where available.

Story	—
BDD	—

REQ0000039 [Requirement]

The Data Inventory Capability must collect the technical and physical inventory of all data elements across the organization.

Story	—
BDD	—

REQ0000042 [Requirement]

The Data Inventory Core Engine must provide the logistical and control framework for individual connections, automatically discovering and maintaining inventory and profiling information within the Data Governance Metadata catalog.

Story	—
BDD	—

REQ0000045 [Fact]

Data Profiling Connectors provide the incremental ability to classify and categorize data based on its content.

Story	—
BDD	—

REQ0000047 [Fact]

Data profiling may occur at a field or record level.

Story	—
BDD	—

REQ0000050 [Requirement]

The Data Management Portal must enable end users to create and maintain inventory metadata manually.

Story	—
BDD	—

7. Non-Functional Requirements

Quality attributes (audit, security, performance, compliance, etc.).

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-202833967-Practical Data- Data Governance At Scale - Data Inventory Capability and Metadata Store.xhtml
Confluence pageId: 202833967
Confluence space: DataModel
Confluence version: 35
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202833967/Practical+Data+Data+Governance+At+Scale+-+Data+Inventory+Capability+and+Metadata+Store
Retrieved at: 2026-02-03T18:25:22.745224

REQ0000048 [Fact]

Data profiling can support identifying that content appears to relate to a person and contains PII information.

Story	—
BDD	—

8. Architecture and Operating Model

Reference architecture, layers, components, and operating model/policy/oversight statements.

Practical Data: Data Governance at Scale - Data Lineage Capability & Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-199491586-Practical Data- Data Governance at Scale - Data Lineage Capability & Metadata Store.xhtml
Confluence pageId: 199491586
Confluence space: DataModel
Confluence version: 42
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/199491586/Practical+Data+Data+Governance+at+Scale+-+Data+Lineage+Capability+Metadata+Store
Retrieved at: 2026-02-03T18:25:24.045921

REQ0000002 [Fact]

Technical Lineage is defined as dependency between a downstream technical attribute and upstream technical attribute dependencies, and can exist within or across systems.

Story	—
BDD	—

REQ0000003 [Fact]

Business Lineage is defined as dependency between logical business definitions of data at the lowest level of the business element hierarchy.

Story	—
BDD	—

REQ0000009 [Fact]

Complex transformations may require introducing a new node element to capture the data transformation rule.

Story	—
BDD	—

Practical Data: Data Governance at Scale - Conventions

Source type: confluence-xhtml
Source: confluence/DataModel-202735632-Practical Data- Data Governance at Scale - Conventions.xhtml
Confluence pageId: 202735632
Confluence space: DataModel
Confluence version: 2
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202735632/Practical+Data+Data+Governance+at+Scale+-+Conventions
Retrieved at: 2026-02-03T18:25:29.764240

REQ0000002 [Requirement]

The document series must employ a context-based naming standard for unique sample IDs to represent sample data governance metadata.

Story	—
BDD	—

REQ0000003 [Fact]

Node ID conventions are specified by NodeType, NodeSubType, and naming convention, with examples provided.

Story	—
BDD	—

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

REQ0000030 [Fact]

An example nodeFQName is crm/database/People/person-details/lastname.

Story	—
BDD	—

REQ0000037 [Fact]

An example nodeFQName is crm/database/People/person-details/lastname.

Story	—
BDD	—

REQ0000038 [Fact]

For INVENTORY/ELEMENT nodes, the nodeSpecification field is present in the model table but its description is not specified in this document.

Story	—
BDD	—

REQ0000041 [Fact]

The Data Inventory Capability includes a Data Inventory Core Engine sub-component.

Story	—
BDD	—

REQ0000043 [Fact]

The Data Inventory Capability includes Data Inventory Connectors.

Story	—
BDD	—

REQ0000044 [Fact]

The Data Inventory Capability includes Data Profiling Connectors.

Story	—
BDD	—

REQ0000046 [Fact]

Data profiling may be implemented as part of a Data Inventory connector or as a separate connector, depending on entitlements and architecture.

Story	—
BDD	—

REQ0000049 [Fact]

The Data Inventory Capability includes a Data Management Portal sub-component.

Story	—
BDD	—

9. Appendices (Diagrams, Navigation, Document Structure)

Facts about diagrams/macros and document navigation that materially affect interpretation.

Practical Data: Data Governance At Scale - Data Inventory Capability and Metadata Store

Source type: confluence-xhtml
Source: confluence/DataModel-202833967-Practical Data- Data Governance At Scale - Data Inventory Capability and Metadata Store.xhtml
Confluence pageId: 202833967
Confluence space: DataModel
Confluence version: 35
Confluence URL: https://maoperatingsystem.atlassian.net/spaces/DataModel/pages/202833967/Practical+Data+Data+Governance+At+Scale+-+Data+Inventory+Capability+and+Metadata+Store
Retrieved at: 2026-02-03T18:25:22.745224

REQ0000051 [Fact]

The document includes an embedded draw.io diagram associated with the Data Inventory Capability section.

Story	—
BDD	—

99. Unassigned (Review Needed)

Catch-all to ensure no requirements are dropped; contains any requirements that did not match earlier sections.

(No items)