Data Cosmos_Auto Data Mapping

AI-Powered Intelligent Data Mapping & Lineage Accelerator

Part of the Coforge Data Cosmos™ Platform | Hypernova™ Portfolio

Coforge Limited | Data HBU

2026

Solution Overview

Data Cosmos_Auto Data Mapping is an AI-powered intelligent accelerator designed to simplify and accelerate complex data modeling and mapping across legacy and modern data landscapes. It serves as an enterprise-grade, AI-driven data mapping and lineage platform engineered to automate, visualize, and govern source-to-target data transformations across heterogeneous systems. Acting as a centralized mapping intelligence layer, Auto Data Mapping leverages AI to analyze schema metadata, suggest high-confidence mappings, and generate comprehensive lineage graphs — making data mapping efficient, accurate, and accessible without requiring extensive manual specification effort.

In modern enterprises, data teams spend 40–50% of their time on manual source-to-target mapping — writing mapping specifications, debugging transformation logic, resolving schema mismatches, and tracing data lineage across systems. Auto Data Mapping eliminates this burden by providing an intelligent mapping fabric that analyzes schema metadata, suggests high-confidence mappings, and provides full lineage traceability — all through a single, unified interface. Whether mapping terabytes of core banking data from Oracle/Teradata to Snowflake on Azure, transforming claims data across legacy mainframe systems, or consolidating patient data across EHR platforms, Auto Data Mapping delivers consistent, governed, and auditable data transformations at enterprise scale.

Key Features

Feature	Description
AI-Powered Auto Mapping	Leverages AI to analyze schema metadata, column names, data types, and business context to automatically suggest accurate source-to-target data mappings with high-confidence recommendations — eliminating manual mapping effort.
Visual Mapping Canvas	Intuitive side-by-side canvas for schema mapping and transformation design. Enables data engineers and architects to visually review, refine, and approve AI-suggested mappings with drag-and-drop simplicity.
End-to-End Data Lineage	Clear lineage graphs ensuring full traceability across all connected systems. Maps data flow from source to target including transformations, joins, and business rules — supporting audit readiness and governance.
Schema Management	Flexible schema management via CSV ingestion or manual input. Enables teams to onboard new sources and targets rapidly, manage schema versions, and maintain a governed catalog of all mapped entities.
AI-Driven Validation	Automated validation engine that cross-checks mapping accuracy across schema types, data formats, and business rules — reducing mapping errors by 40%.
Collaboration & Export	Multi-user workspace enabling cross-functional collaboration between engineering, analytics, and governance teams with exportable mapping artifacts for downstream ETL/ELT pipelines.

4-Stage Intelligent Pipeline

Stage	Role	What It Does
AI-Powered Auto Mapping	Intelligent Schema Analysis & Mapping	Analyzes source and target schema metadata, column names, data types, and business context. Generates high-confidence source-to-target mapping recommendations automatically — eliminating the need for manual mapping specification documents.
Visual Mapping Canvas	Interactive Design & Refinement	Provides an intuitive side-by-side canvas where data engineers and architects visually review, refine, and approve AI-suggested mappings. Supports drag-and-drop simplicity, transformation rule definition, and real-time mapping validation.
End-to-End Data Lineage	Traceability & Governance	Generates clear lineage graphs mapping data flow from source to target including all transformations, joins, filters, and business rules. Ensures audit readiness, regulatory compliance, and full governance traceability across all connected systems.
Schema Management Layer	Catalog & Version Control	Supports flexible schema onboarding via CSV ingestion or manual input. Manages schema versions, maintains a governed catalog of all mapped entities, and enables rapid onboarding of new sources and targets.

How Auto Data Mapping Works

1. **Discover** — Connect to source and target systems, explore schemas, metadata, and sample data across the enterprise landscape
2. **Ingest Schemas** — Import source and target schemas via CSV upload, API connectors, or manual definition for mapping setup
3. **AI Analysis** — AI engine analyzes schema metadata, column names, data types, patterns, and business context across source-target pairs
4. **Auto Mapping** — Generate high-confidence source-to-target mapping recommendations automatically based on AI analysis
5. **Visual Review** — Data engineers review, refine, and approve AI-suggested mappings on the interactive visual mapping canvas
6. **Transformation Design** — Define transformation rules, business logic, joins, and filters within the visual canvas for each mapping
7. **Lineage Generation** — Auto-generate end-to-end data lineage graphs showing complete data flow from source to target with all transformations
8. **Validation** — AI-driven validation cross-checks mapping accuracy, identifies conflicts, and flags potential data quality issues
9. **Export & Handoff** — Export mapping specifications, lineage documentation, and transformation logic for downstream ETL/ELT pipeline development
10. **Govern** — Maintain versioned mapping catalog, audit trails, and compliance documentation for regulatory and governance oversight

Industry Trends and where it can be leveraged

Banking & Financial Services

Use Case	Example
Regulatory Data Migration Mapping	A global bank maps 500+ tables from legacy core banking systems (Temenos, Finacle, Oracle) to a centralized Azure Synapse/Snowflake warehouse for BCBS 239 regulatory reporting. Auto Data Mapping's AI engine analyzes complex legacy schemas and generates high-confidence mapping recommendations, while lineage graphs provide end-to-end audit trails required by regulators for data provenance verification.
Payments Transformation Mapping	A retail bank maps payment transaction data across multiple payment rails (SWIFT, ACH, real-time payments) into a unified analytics platform. The visual canvas enables payment architects to validate domain-specific transformation logic while AI-driven validation ensures mapping accuracy across diverse payment message formats.

Insurance

Use Case	Example
Claims Data Platform Mapping	A multi-line insurer maps claims data from legacy Guidewire ClaimCenter and Duck Creek systems to Snowflake for a unified claims analytics platform. Auto Data Mapping handles schema mapping across disparate claim formats (auto, property, liability), with AI-driven validation ensuring actuarial logic fidelity and lineage graphs providing complete traceability for compliance teams.
Underwriting Data Integration Mapping	An underwriting team maps third-party risk data feeds, policy administration systems, and actuarial models into a centralized underwriting data mart. Schema management enables rapid onboarding of new data providers while maintaining governed version control across all mapped entities.

Travel, Transportation & Hospitality

Use Case	Example
Reservation & Revenue Data Mapping	A global airline maps real-time booking data from Amadeus/Sabre GDS, loyalty program transactions, and ancillary revenue streams into a Databricks Lakehouse. Auto Data Mapping's visual canvas enables revenue management teams to validate complex domain-specific mapping logic across fragmented airline systems, with lineage graphs ensuring full traceability for revenue audit.
Customer 360 Data Mapping	A hotel chain maps guest profile data from PMS (Opera), CRM (Salesforce), loyalty platforms, and OTA channels into a centralized Customer 360 on Azure Data Lake. AI-powered mapping handles identity resolution across systems with different guest ID formats, while schema management maintains governed mapping versions across all source systems.

Healthcare & Life Sciences

Use Case	Example
Clinical Data Lake Mapping	A healthcare provider maps EMR/EHR data (Epic, Cerner) along with lab results, imaging metadata, and claims data into a HIPAA-compliant Azure Data Lake. End-to-end lineage ensures full traceability for PHI/PII data flows, while schema management maintains HIPAA-compliant data handling standards across all mapped entities.
Clinical Trial Data Mapping	A pharma company maps clinical trial data from multiple CROs, CTMS platforms, and EDC systems (Medidata Rave, Oracle InForm) into a unified research data platform. Auto Data Mapping automates schema analysis across diverse file formats (SAS, CSV, XML), enabling biostatisticians to access harmonized trial data 40% faster.

Cross-Industry Use Cases

Use Case	Description
Source-to-Target Migration Mapping	Automate mapping specifications for large-scale data migrations from legacy systems (Oracle, Teradata, SQL Server) to modern cloud platforms (Snowflake, Databricks, Azure Synapse) with AI-driven recommendations.
Data Warehouse Modernization	Map complex legacy warehouse schemas to modern lakehouse architectures with full lineage traceability and governed transformation documentation.
Regulatory & Compliance Mapping	Generate audit-ready lineage graphs and mapping documentation for regulatory frameworks (GDPR, HIPAA, SOX, BCBS 239) with complete data provenance trails.
M&A Data Integration	Accelerate post-merger data integration by rapidly mapping and reconciling overlapping schemas, business entities, and data definitions across acquired systems.
Multi-Cloud Data Platform Mapping	Map data flows across AWS, Azure, and GCP environments with centralized lineage visibility and governed mapping artifacts.

Target Audience

- Enterprise Data Engineering & Platform Teams
- Data Architects and Cloud Migration Leads
- Data Modelers and Integration Specialists
- Data Quality & Governance Teams
- Analytics Leaders and Business Intelligence Teams

Azure Compatibility

Auto Data Mapping integrates natively with Azure OpenAI Service for AI-powered mapping intelligence, Azure Kubernetes Service (AKS) for scalable deployment, Azure Data Lake Storage for schema and artifact storage, and Azure Synapse Analytics as a target platform. It also integrates with Azure Data Factory, Snowflake on Azure, Databricks on Azure, and Microsoft Purview for governance. It is part of the Coforge Data Cosmos Hypernova portfolio for greenfield data programs and advanced analytics, fully designed to leverage Azure-native services for scalable, secure, and governed enterprise data mapping and lineage.