

INTRODUCING

# Revolutio DataHub

## The governed enterprise data operating layer

Unified ingestion. Curated data products. Trusted analytics.



**Revolutio DataHub** Data Steward

Search data products, assets, pipelines...

- Home
- Ingestion
- Preparation
- Data Products
- Governance
- Catalog
- Lineage
- Quality
- Observability
- Settings

### Platform overview

<b>Pipelines</b> 128 <small>Active pipelines</small>	<b>Data Products</b> 86 <small>Published</small>	<b>Datasets</b> 4.2 TB <small>Curated data</small>	<b>Quality Score</b> 98% <small>Overall score</small>
--	--	--	---

### Recent data products

Name	Domain	Status	Quality	Updated
Customer 360	Customer	Published	98%	2h ago
Order Insights	Sales	Published	97%	3h ago
Product Master	Product	Published	98%	4h ago
Supply Chain View	Supply Chain	Published	96%	5h ago
Finance Snapshot	Finance	Published	97%	6h ago

### Pipeline health

Healthy 108 | Warning 12 | Failed 8

Last 7 days



### Ingest & unify

Connect any source and unify batch, streaming and real-time data at scale.



### Prepare & enrich

Transform, cleanse and enrich data to create high-quality, curated data products.



### Govern with control

Policy-driven governance, lineage, quality and access control built in.



### Consume with confidence

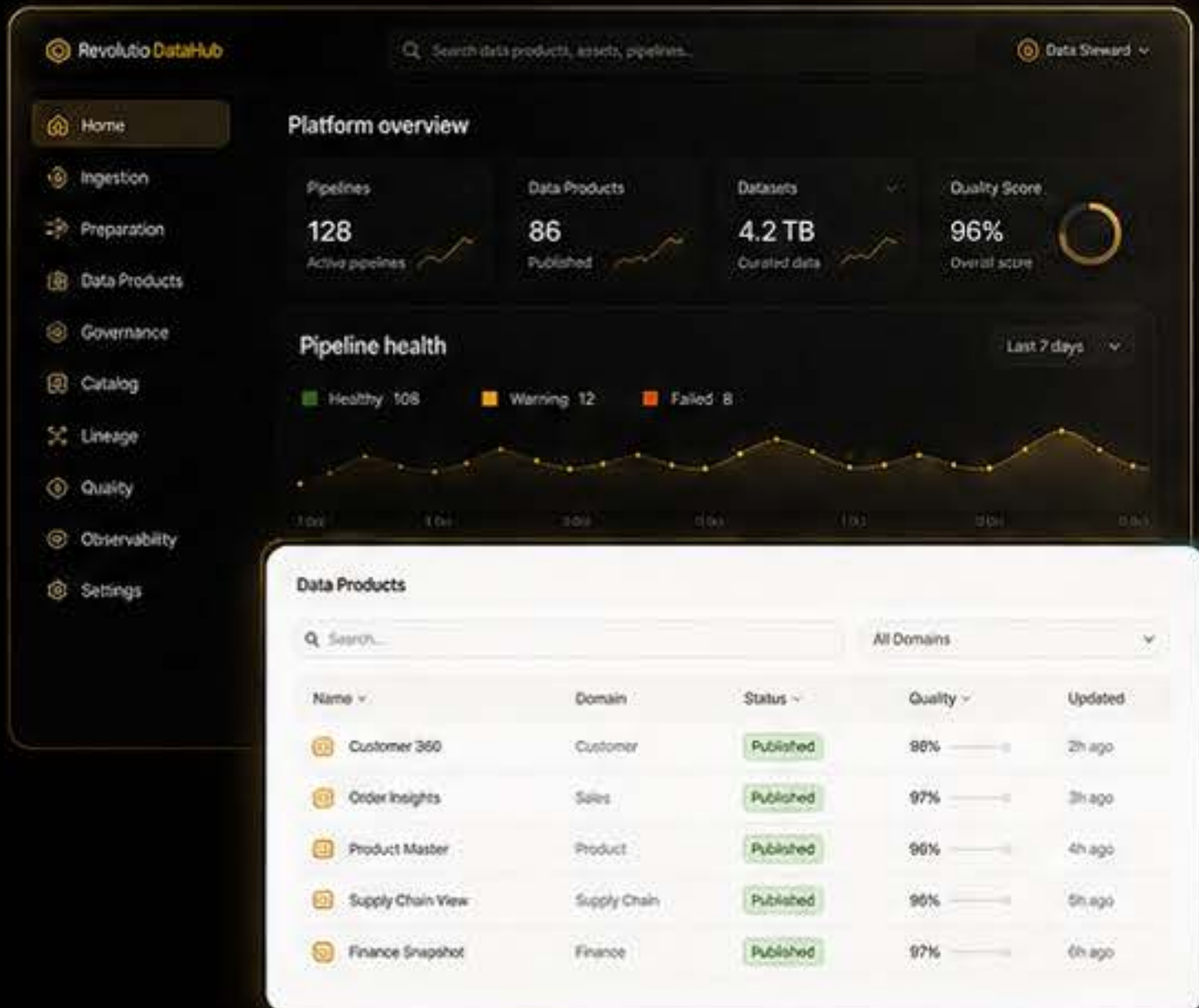
Discover, trust and use governed data and analytics across the enterprise.

WHY

# Revolutio DataHub

A modern governed data platform for enterprises that need speed, trust and control.

Revolutio DataHub is a next-generation enterprise data platform that unifies ingestion, transformation, governance and analytical consumption in a single, governed operating layer. Built for scale and designed for trust, it empowers data teams to deliver reliable, high-quality data and actionable insights—faster.



## Built for enterprise-scale data operations



### Data Ingestion & Unification

Connect any source and ingest batch, streaming and real-time data with schema capture and change tracking.



### Data Preparation & Enrichment

Profile, cleanse and enrich data with low-code data flows and reusable business rules.



### Transformation Lifecycle Management

Design, version and orchestrate transformations with built-in testing, lineage and deployment controls.



### Governance, Security & Compliance

Policy-driven governance, cataloging, data quality, access control and audit logging built in.



### Data Lineage & Traceability

End-to-end lineage across pipelines, datasets and data products for full impact analysis.



### Analytics & Consumption

Deliver governed data products that power dashboards, data apps and AI/ML with one-click activation.

## Built different. Built for results.



### No source mutation

Operate without altering source systems. Preserve integrity at every step.



### Live + incremental ingestion

Capture changes in real time for always-fresh, decision-ready data.



### Row-level security

Fine-grained access controls ensure the right data to the right people.



### One-click BI activation

Publish trusted data products and activate directly in your BI tools.



Designed for organizations evaluating alternatives to Palantir Foundry, Microsoft Fabric and Databricks.

DATA INGESTION & UNIFICATION

# Ingest. Unify. One Databucket.

Ingest data from multiple heterogeneous sources into a single, trusted Databucket ready for enterprise consumption.



The screenshot shows the Revolutio DataHub interface. At the top, there's a search bar and a 'Data Steward' dropdown. A sidebar on the left lists navigation options: Home, Ingestion, Preparation, Data Products, Governance, Catalog, Lineage, Quality, Observability, and Settings. The main area displays a 'Pipelines' overview with summary cards for 128 Active pipelines, 86 Running now, 12 Failed, and 4.2 TB Ingested (7d). Below this is a table of pipelines with columns for Pipeline name, Source, Type, Status, Frequency, Last run, and Records. The 'Pipeline details' section below shows a flow from 'Source PostgreSQL' (Tables: 12) through a 'Transform & Map' step (Schema mapping, Type casting, Validation, Deduplication) to 'Databucket Customer 360' (Incremental upsert, No write-back). A 'Schedule' panel on the right shows 'Every 5 minutes', 'Last run: 1 min ago', and 'Status: Running'.

**Multiple heterogeneous data sources**

Connect to databases, applications, files, APIs, streaming platforms, cloud apps and warehouses.

**Incremental & live data ingestion**

Capture changes in real time or in near real time with incremental, checkpointed loads.

**No manual refresh cycles**

Automated scheduling and change detection keep your data always fresh.

**Source-safe by design**

Preserve source integrity with read-only ingestion and no write-back operations.

**Single trusted Databucket**

All sources unified into one governed, consumable layer for analytics and AI.

**KEY OUTCOMES**

Always-fresh data

Source-safe ingestion

Unified consumption

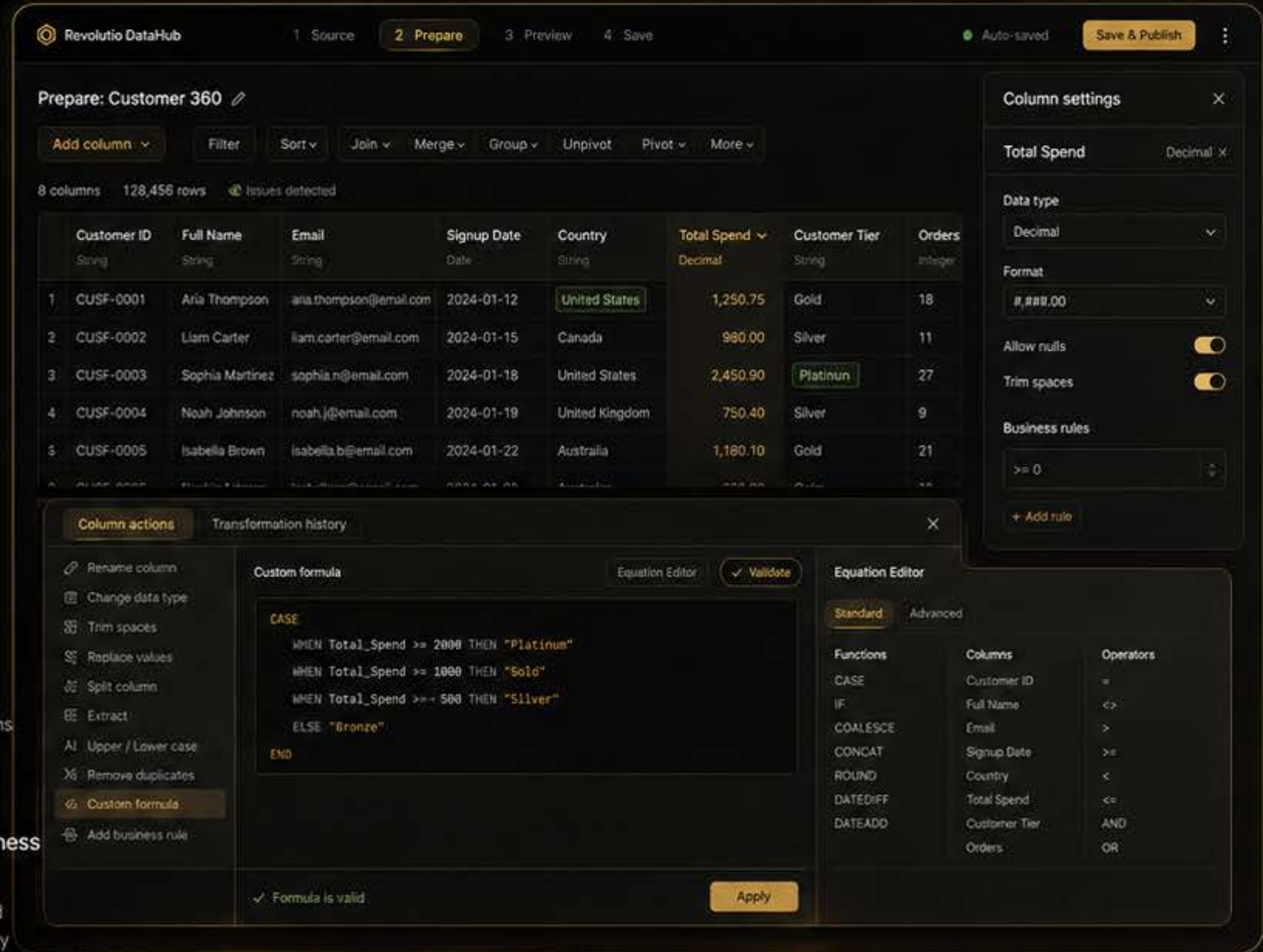
Operational simplicity

HOW YOU PREPARE


# Prepare. Enrich. Curate.


Transform raw data into trusted, analysis-ready datasets—without touching the source.

Revolutio DataHub empowers business and data teams to cleanse, shape and enrich data with powerful, intuitive tools. Apply transformations, build logic, enforce rules and combine datasets—fast, safely and at scale.





-  Low-code preparation
-  Reusable logic
-  Fast joins & merges

 **Row-level & column-level cleansing**  
Trim, standardize, deduplicate and correct data across rows and columns with powerful built-in actions.

 **Visual Equation Editor**  
Build standard or advanced calculations using a visual editor with functions, operators and column references.

 **Derived & custom columns**  
Create calculated fields, flags and segments using intuitive logic and reusable expressions.

 **Column-level formatting, business rules & measures**  
Enforce formats, apply validations and create measures to ensure data quality and consistency.

 **Seamless data type changes**  
Convert, cast and harmonize data types effortlessly to fit downstream needs.

 **High-performance merges & joins**  
Join and merge datasets of any size in seconds—no code, no compromises.



TRANSFORMATION LIFECYCLE MANAGEMENT

# Build. Version. Evolve. Transform with control.

From raw ingestion to trusted output—a fully governed, versioned transformation journey you can trust.

Design, iterate and publish transformations with full visibility, impact awareness and end-to-end governance.



**Step-wise, versioned pipeline**  
Model transformations as modular steps with versioning, rollback and impact tracking.

**Full transformation visibility**  
End-to-end lineage, step dependencies and data flow insights across the entire journey.

**Edit, remove or reorder with confidence**  
Modify steps in any order without breaking downstream flows or rework.

**Preview at every stage**  
Preview intermediate and final datasets to validate results before publishing.

Transformation Pipeline
Customer 360 – v2.1 (Draft) Draft
Owner: Data Engineering
Last modified: May 19, 2025 10:24 AM
Version history: v2.1 • v2.0 • v1.3 Publish

- 1 **Ingest Source**  
customer\_raw
- 2 **Filter & Cleansing**  
Remove nulls, duplicates
- 3 **Enrich & Standardize**  
Add reference data
- 4 **Apply Business Rules**  
Derive segments
- 5 **Aggregation & Metrics**  
Compute KPIs
- 6 **Final Output**  
customer\_360\_curated

[+ Add Step](#)

### Step 1: Ingest Source

Configuration Preview Lineage Impact

Preview (customers\_raw) 5 columns • 12,842 rows

customer_id	name	email	signup_date	country
10001	Jane Cooper	jane.cooper@acme.com	2024-12-01	US
10002	Albert Flores	albert.flores@acme.com	2024-12-02	UK
10003	Arlene McCoy	arlene.mccoy@acme.com	2024-12-03	CA
10004	Wade Warren	wade.warren@acme.com	2024-12-04	US
10005	Dery Bhumish	dev.bh@acme.com	2024-12-05	IN

Columns: 5

Rows: 12,842

Data quality: 96%

Last refreshed: 2 min ago

#### Pipeline details

Name: Customer 360 Transformation

Description: Builds the curated Customer 360 dataset from raw sources with standardization, enrichment and business rules.

Tags: customer, curated, marketing

Schedule: Daily at 02:00 AM

---

#### Version

v2.1 (Draft)

By Data Engineering

May 19, 2025 10:24 AM

[View all versions](#)

**Change safely. Iterate faster. Govern better.**  
Versioned transformations with full visibility and control—so your data stays trusted at every step.

GOVERNANCE, SECURITY & COMPLIANCE

# Govern trusted data operations

Security and control by design.

Revlutio DataHub embeds governance, security, and compliance into every layer of the data lifecycle. Define who can access what, how changes are approved, and how every action is recorded—so you stay in control and always audit ready.

**Approval workflows**  
Maker-checker approval flows for access requests, rule changes and transformations.

**Hierarchical security**  
Define rules using natural hierarchies like Region → Country → Branch for precise data access.

**RBAC**  
Role-based access control with fine-grained permissions across assets, pipelines and data operations.

**Auditability**  
Complete audit trails of data changes, approvals, access rules and execution history—always traceable.

Trusted operations | Fine-grained access | Runtime enforcement | Regulatory-ready traceability

LINEAGE TO INSIGHT

# End-to-end lineage. Trusted data in action.

Trace every transformation. Understand every impact.  
Activate governed data for trusted analytics—instantly.



- End-to-end lineage from source to consumption
- Track movement across transformations, enrichments and security layers
- Identify downstream impact when source data or logic changes
- Support regulatory reporting, audits and impact analysis
- One-click integration with Revolutio BI
- Publish Databuckets as trusted, governed analytical datasets
- Automatically enforce row-level and role-based security cutting consumption
- Support real-time or near-real-time analytics through live data updates

**Trace every step**  
Complete visibility from source to dashboard with full lineage and audit trails.

**Know downstream impact**  
Instantly discover what's affected by changes to data, logic or policies across the pipeline.

**Activate governed analytics**  
Publish trusted Databuckets and one-click activate in Revolutio BI to drive business insights.

**Consume in real time**  
Always up-to-date insights with live data updates and performance built for decision speed.

## From governed Databuckets to powerful insights

Activate curated data in Revolutio BI for trusted, self-serve analytics.






**One-click activation to Revolutio BI**  
Secure. Governed. Always up to date.

[Open in Revolutio BI →](#)

# Built to scale. Designed to win.

Performance, governance and trust—  
at enterprise scale.

## PERFORMANCE & SCALABILITY

- 
**Execute transformations and merges at high speed**  
 Vectorized pipelines and optimized compute deliver faster processing with lower latency.
- 
**Handle live data updates automatically**  
 Incremental ingestion, change capture and orchestration keep data fresh without manual intervention.
- 
**Scale Databuckets seamlessly**  
 Elastic architecture grows with your data volume, users and transformations—without replatforming.
- 
**Optimize performance while maintaining control**  
 Built-in governance, security and lineage ensure speed never overcomes trust.

## WHY REVOLUTIO DATAHUB



**Unified platform**  
One governed layer for ingestion, preparation, governance and consumption.



**Governed by design**  
Policy-driven controls, lineage and observability built in—so trust never becomes a bottleneck.



**Low-code productivity**  
Intuitive UI, reusable components and automation empower teams to deliver more, faster.



**Trusted analytical activation**  
Curated, high-quality data products that are ready for BI, AI/ML and advanced analytics.



**Enterprise scale**  
Elastic, resilient architecture that grows with your business—securely and reliably.



**Launch a governed data foundation for the enterprise**

Revolutio DataHub unifies ingestion, preparation, control and consumption into one trusted operating layer.

Ready to build your advantage?



## SCALE WITHOUT LIMITS

- Data**  
Scale to petabytes of data
- Users**  
Support thousands of users
- Transformations**  
Run millions of operations
- Performance**  
Consistent speed and reliability

The modern, governed data layer for enterprises evaluating Palantir Foundry, Microsoft Fabric and Databricks.