

Build AI Governance from the Ground Up with CData Sync



Stanley Liu

Associate Technical Product Marketer, CData
stanleyl@cdata.com



Cameron Leblanc

Technology Evangelist, CData
cameronl@cdata.com

Agenda

- The AI governance challenge
- Building the foundation
- Auditable and compliant pipelines demo
- Customer spotlight
- Key takeaways
- Q & A



CData Software

Based in Chapel Hill, NC, with offices in Europe, Japan, and India, CData is the leader in real-time data connectivity software for enterprises, developers, and data scientists.

400+

EMPLOYEES

10,000+

ACTIVE CUSTOMERS

140+

EMBEDDED
PARTNERS



CData Sync

Easy to use, real-time modern data pipelines

**Automated continuous data replication
from any source to any data store**

- Enterprise-grade connectors with unmatched speed and depth
- Deploy anywhere — on-prem, private cloud or SaaS
- Clustering – parallel processing across Sync instances
- Robust metadata and change data capture (CDC) capabilities; real-time data synchronization
- Transform data using dbt or SQL
- Predictable, scalable pricing based on connections (not rows)

Financial Planning and Analytics Scenario

Business challenge

- A global organization needs to consolidate financial, operational, and budgeting data scattered across ERP, HR, CRM, and planning systems into an on-premises data warehouse
- Goal: Enable real-time forecasting, variance analysis, and AI-assisted financial planning and reporting using complete, timely, and trusted data

Why AI Governance Matters

- Both predictive and generative AI require governed, up-to-date data to produce reliable forecasts, insights, and scenario simulations
- Many organizations must keep sensitive financial data on-premises so local AI agents can operate securely without exposing regulated information

How Sync helps

- Unifies on-prem and cloud financial, operational, and planning data sources for analytics and AI
- Keeps financial and operational data continuously updated with Change Data Capture (CDC) for real-time forecasting
- Supports both structured forecasting models (predictive AI) and unstructured FP&A knowledge retrieval or planning assistants (RAG/assistant use cases)

The foundation: Governed, unified data pipelines

The boring work that makes AI successful

Unify everything

Unify data from every source, cloud, and legacy system

Keep it fresh

Preserve data quality and integrity across hybrid and multi-cloud environments

Automate compliance

Data pipelines are made with built-in traceability and observability

Combine old with new

Connect to on-prem/legacy systems without the need to rip and replace

Demo

Governance that keeps AI reliable

Every AI use case depends on consistent, curated, and well-governed data.

Scheduling & refresh – Control cadence and cost by defining when and how data moves

Change Data Capture (CDC) – Keep every pipeline current with minimal overhead

Scope & selection – Replicate only what matters (specific tables, columns, or records)

Schema & transformation rules – Manage changes without breaking downstream models

Demo

Effortless Data Compliance

Data pipelines are built with **built-in traceability** and **full observability**, ensuring continuous compliance without manual effort.

- Every data movement is logged for clear, auditable traceability
- Pipeline steps are fully recorded for compliance documentation
- Monitoring job status ensures compliant data delivery
- Automatic error handling prevents unapproved or incomplete transfers



Demo

No-code pipeline automation powers patient insights

Empower Services reduced AI pipeline runtime by 91%



Empower's legacy pipelines—using SQL log shipping and SSIS—were brittle, slow, and required manual table mapping, leading to delays and wasted engineering effort. A **12-hour lag** made the data unreliable for timely analytics.

91% reduction in pipeline runtime

From 90 minutes to 8 minutes

6x more frequent Snowflake updates

Move 3x more tables for fresher, more complete data

Modern data movement for Cortex AI

- **Improved reliability and reduced compute load** with incremental replication, tuning, and flexible scheduling
- Enabled AI-driven use cases through **Snowflake + Cortex** integration
- Standardized CDC framework with **schema change & error handling**
- Accelerated Snowflake ROI — realizing value up to **10x faster** without consumption-based pricing

AI success starts with governed, transparent data movement

Any data source

Operational,
SaaS, and legacy



Governed, continuously
updated pipelines

Any AI use case

AI-driven analytics, vector
search, and copilots



Google BigQuery



Q&A



More About Sync

Try Sync Free for 30 Days:
www.cdata.com/sync/trial/

Contact: sales@cdata.com