

DATASHEET

Pentaho™ Data Integration - Metadata Injection

Accelerate Complex Data-Onboarding Projects

Modern Data-Onboarding Challenges

Today's data-onboarding projects require more than just connecting to or loading data. Rather, they involve managing a changing array of data sources, establishing repeatable processes at scale, and maintaining control and governance. Whether an organization is implementing an ongoing process for ingesting hundreds of data sources into Hadoop or enabling business users to upload diverse data without IT assistance, onboarding projects tend to create major obstacles:

ETL User Challenges

- Repetitive manual design.
- Time-consuming development.
- Manual error risks.

Business Challenges

- Monopolizing IT resources.
- Putting deadlines at risk.
- Opportunity costs.

The Solution Lies in the Metadata

Hitachi Vantara's unique metadata injection capability allows one transformation to become many, drastically boosting productivity and reducing development time. Pentaho Data Integration, from Hitachi Vantara, is able to access source metadata at run time, which can be injected, on the fly, into a transformation template to control complex logic.

The "instructions" derived from field names, types, lengths and other metadata can dynamically generate the actual transformations, drastically reducing time spent designing transformations.

Further, the metadata can be detected from the data source itself, from separate files, or end user input via API calls. The metadata injection process can be scaled to cover hundreds or even thousands of data sources (see Figure 1).

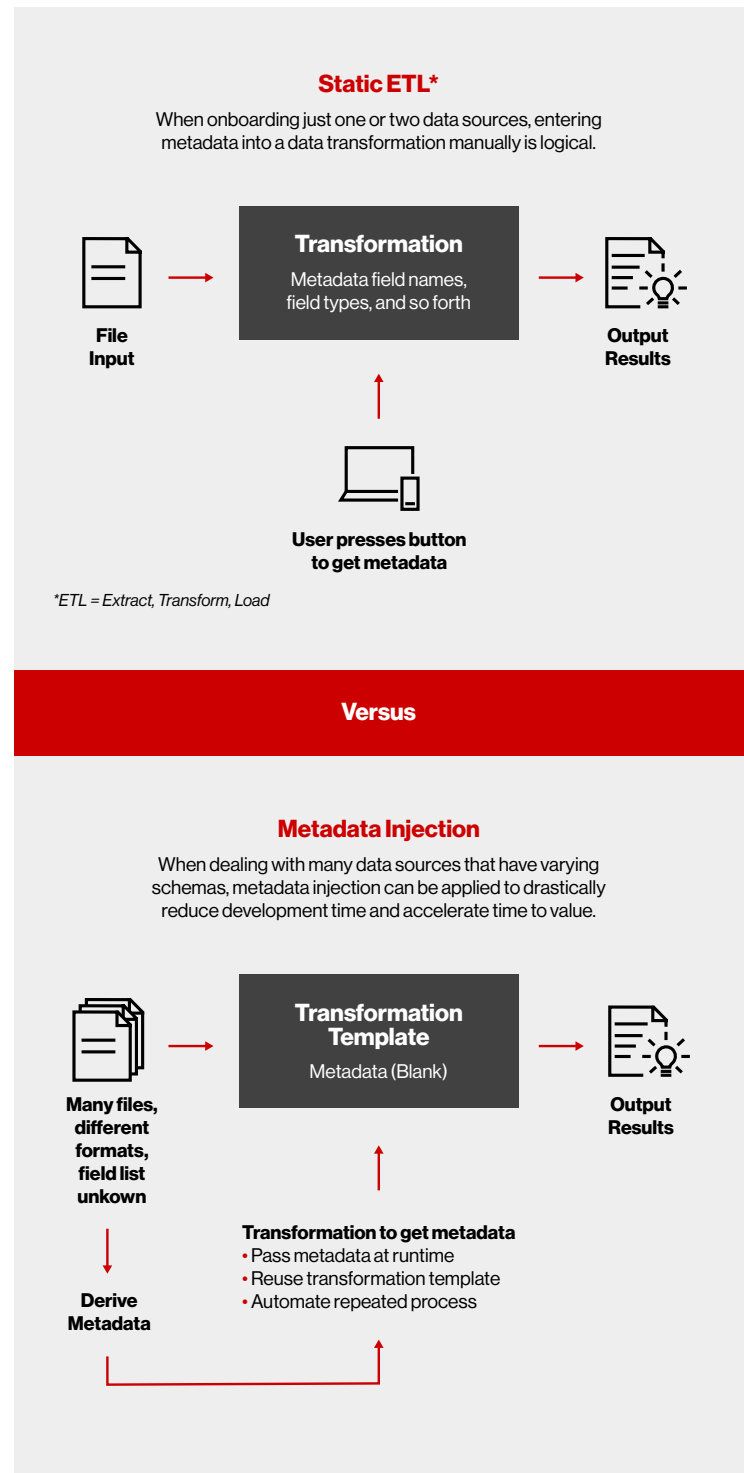


Fig. 1 Compare static ETL with metadata injection for data-onboarding.

Use Case: Kingland Systems

Challenge

- Help Fortune 500 banks meet compliance and risk management requirements by better understanding their data.
- Alleviate manual processes required to onboard customer data into the Kingland application, which provides a single view into data quality.

Pentaho Solution

- Empower Kingland customers to onboard data in a self-service fashion.
- Help create a “one stop shop” for bank data quality, using metadata injection to help automate the process of loading and parsing customer data.

Solution Blueprint: Self-Service Data Onboarding

This use case design pattern leverages Pentaho products and services expertise (see Figure 2). Metadata injection is drawn upon to drastically reduce time needed to onboard customer data into a central system, which could either be a relational environment or Hadoop.

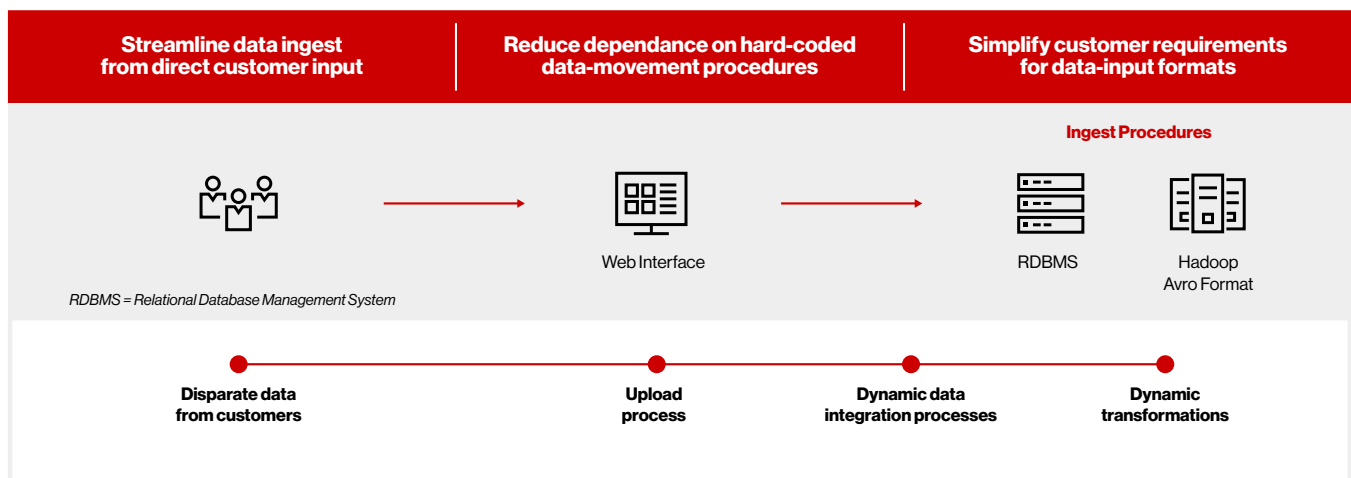


Fig. 2 Solution Blueprint

ABOUT HITACHI VANTARA

Hitachi Vantara, a wholly-owned subsidiary of Hitachi Ltd., delivers the intelligent data platforms, infrastructure systems, and digital expertise that supports more than 80% of the Fortune 100. To learn how Hitachi Vantara turns businesses from data-rich to data-driven through agile digital processes, products, and experiences, visit hitachivantara.com.



Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA
hitachivantara.com | community.hitachivantara.com

Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526
hitachivantara.com/contact

© Hitachi Vantara LLC 2023. All Rights Reserved. HITACHI and Pentaho are trademarks or registered trademarks of Hitachi, Ltd. All other trademarks, service marks and company names are properties of their respective owners.

HV-CBE-DS-Pentaho-Data-Integration-Metadata-Injection-7Jul23-B