pentaho
A Hitachi Group Company

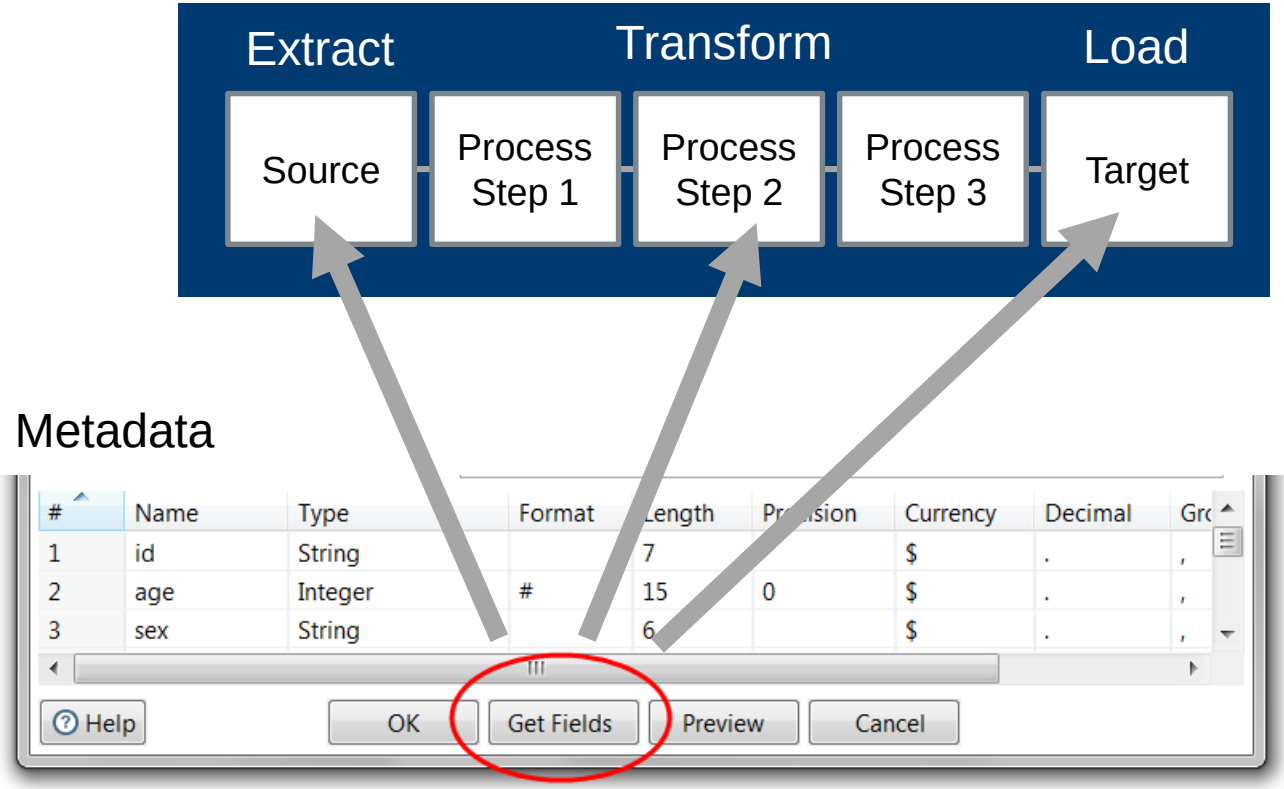
ETL Metadata Injection

A template-based approach to dynamic data integration

Traditional ETL – Hardcoded Metadata

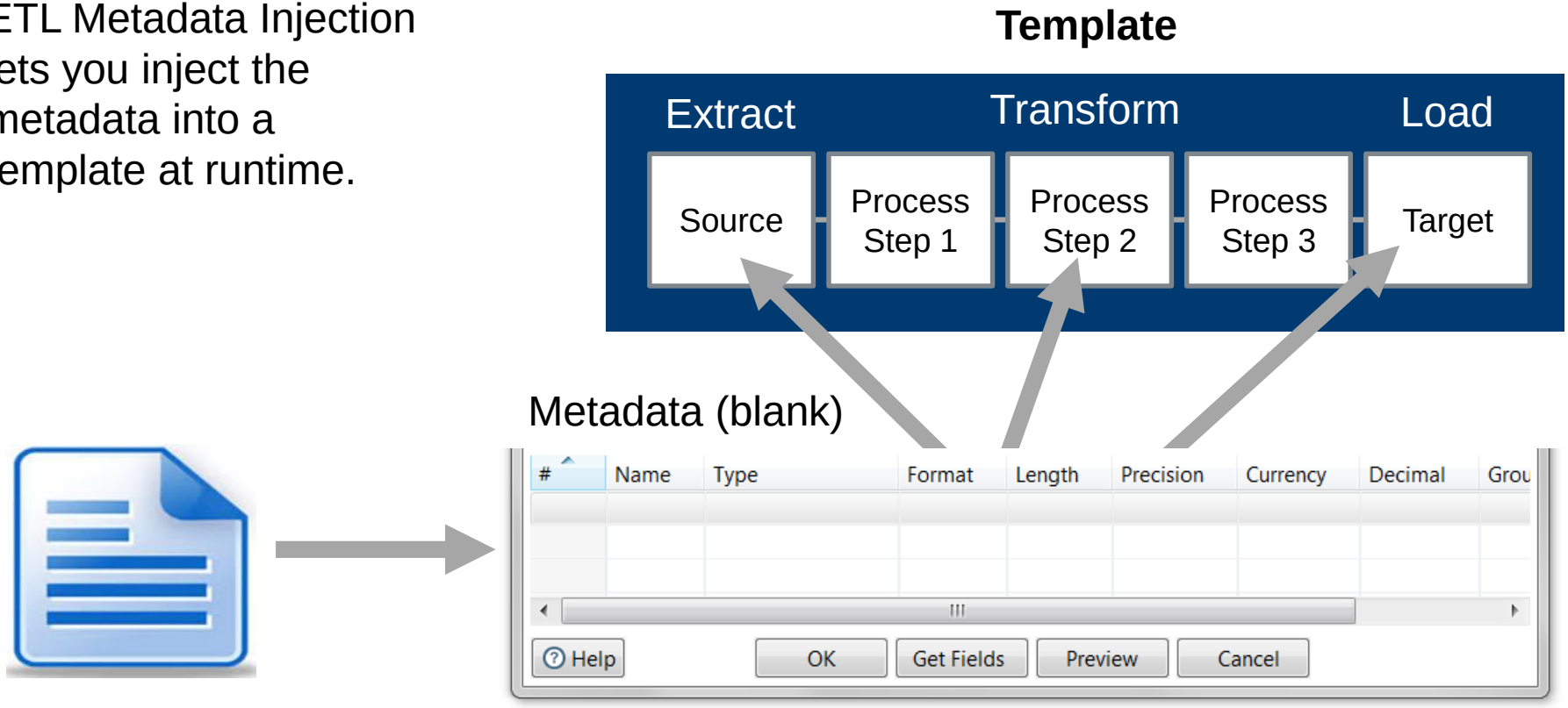
Metadata details (fields, datatypes, etc.) are required for various steps within a transformation: sources, targets, and/or transformation steps.

Legacy ETL tools require you to hardcode the metadata at development time.



Dynamic ETL

ETL Metadata Injection lets you inject the metadata into a template at runtime.



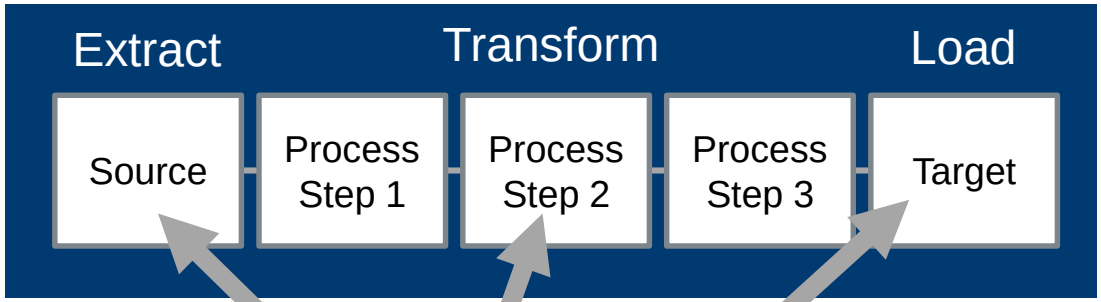
Use Case 1 – Scalability / Reuse

Same workflow, many different files/tables, etc.

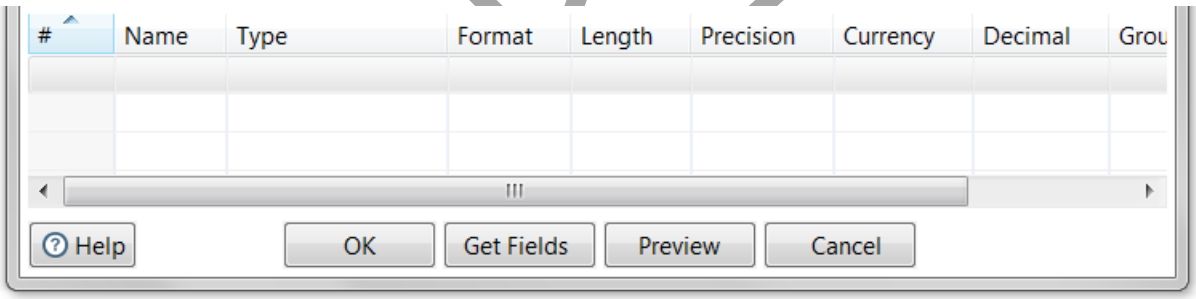
Maintain metadata in a list/table and reuse a single workflow template.

Example:
migrate 1,500 tables

Template



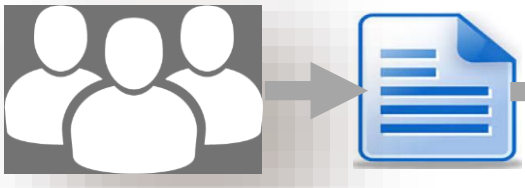
Metadata (blank)



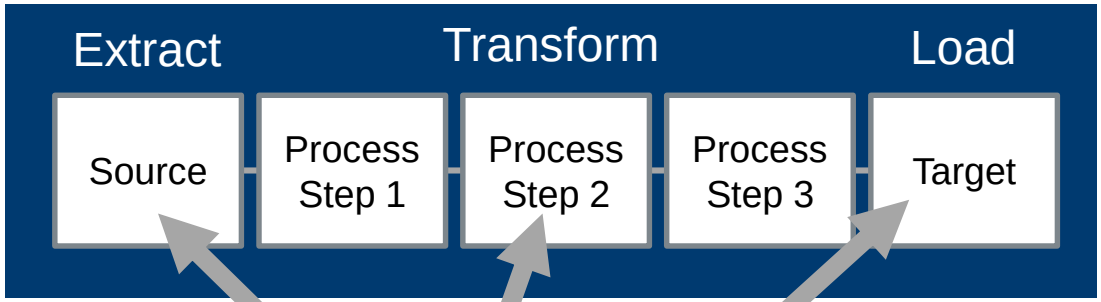
Use Case 2 – Self-service

Allow user/customer to enter metadata in a simple web form

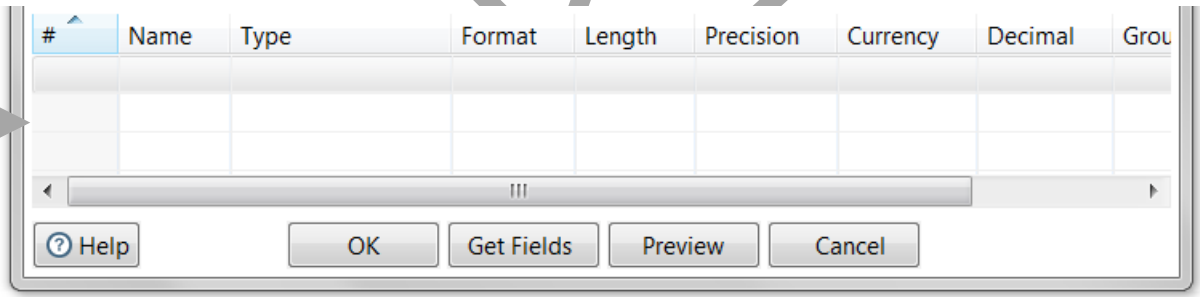
Example:
select fields for a template to pull data from Hadoop and build an on-demand data mart



Template



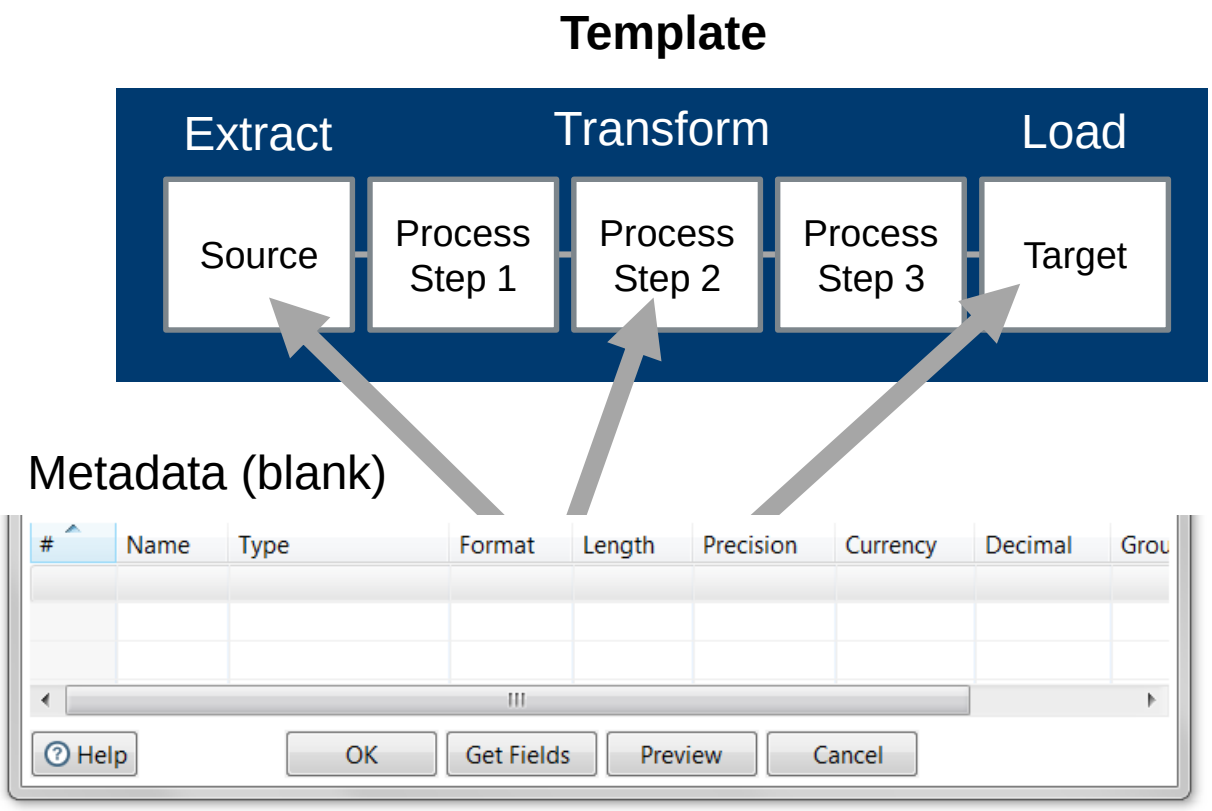
Metadata (blank)



Use Case 3 – Auto-Discovery

Parse out metadata dynamically at runtime.

Example:
Dynamically parse messages of varying formats

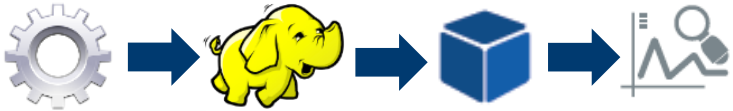


DRY Principle – Don't Repeat Yourself

Use a Templated Approach

Deutsche Bank

Large Oil & Gas Co.



KINGLAND SYSTEMS

Major Professional Services Firm

Use Cases

Scalability: simplified data on-boarding & management

Auto-Discovery: dynamic parsing of log files for cyber-security

Self-service: customer on-boarding

Scalability: large data migration