leaplogic

Automated workload transformation from Informatica to Databricks

VIDEO TRANSCRIPT

LeapLogic eliminates the slow, error-prone, expensive practices of manual migration and automates end-to-end enterprise data warehouse transformation. We bring all the pieces of the puzzle together so that your journey to a cloud-native stack is fast, affordable, and risk-free.

Here's a demo of how LeapLogic simplifies the migration of Informatica legacy workloads to Databricks Lakehouse.

It first assesses the existing Informatica ETL scripts and provides comprehensive complexity analysis and end-to-end data and process lineage. This helps offload workloads to Databricks Lakehouse in a phased manner.

Here is a summary of the assessment insights. We uploaded 25 files, which contained 299 mappings and 1238 transformations. Overall, the system was able to analyze 91 percent of the given ETL scripts successfully. This means it will automate up to 91 percent of Informatica code transformation to the Databricks equivalent.

For these ETL scripts, LeapLogic identifies workflows, mappings and transformations.

Based on various parameters, it calculates the level of complexity for each file. This helps identify the nature of workloads and level of transformation effort required. It also details out end-to-end data and process lineage, impacted entities, missing tables and more.

A comprehensive, graphical lineage showcases complex interdependencies between different workloads.

On hover, it highlights the chain of dependencies for that specific file. Let's see more details.

It highlights the used and impacted tables for the Informatica XML.

A detailed lineage report can also be downloaded for offline use.

This page lists source, lookup, and target entities.

This provides extensive downloadable reports with comprehensive insights.

The reports are downloaded in the CSV format.

Next, the assessed ETL scripts are ready to move to the Transform step, which automatically converts the Informatica flows to Databricks. LeapLogic uses an intelligent grammar engine to convert all mappings and transformations to a Databricks equivalent.

Let's use the drag-and-drop based pipeline interface, where we have an executed Informatica ETL conversion pipeline. One of the files shows the conversion of complex Informatica workflows and mappings.

Let's open this in edit mode.

See the Informatica conversion pipeline where we have configured the data sources, selected the Informatica XML that we want to transform and the required target.

We are selecting Databricks Lakehouse as the target. We will configure and execute this pipeline to transform the ETL workloads using LeapLogic's intelligent grammar engine.

Now let's see the detailed reports.

This shows the converted transformations which contained source qualifiers, Informatica expressions, lookups, filters, and more.

We can see the transformation-level report for each mapping.

Let's download the converted package.

This package contains the Databricks-equivalent PySpark code.

Next, we log in to the Databricks portal and import the converted PySpark code. Let's take a look.

For the Databricks equivalent code, we first establish connection with the data source, and then execute the transformations. In the end, we load the final output into the required table.

Here is the converted code snippet with transformation details. In this code, first the source qualifier query is converted, then, an expression, followed by join, and others.

We can now execute the converted code. This concludes the Transformation step.

Let's see the list of jobs.

Filter the job created by you and execute.

LeapLogic's holistic approach further includes validation and operationalization of migrated workloads on Databricks Lakehouse.

For over a decade, we've helped Fortune 100 companies take data-driven decisions with absolute confidence and drive powerful business outcomes they'd never even imagined.

We can guide you through this process, starting with an assessment and proof of concept to build the foundation for a successful project. Our support model is flexible based on your needs, ranging from providing transformed and validated code up through end-to-end transformation services. We're here to accelerate your modernization journey.