ETL TOOLS

1. Talend

Talend is an open-source data integration and management platform that enables big data ingestion, transformation and mapping at the enterprise level.

**Pros**: **Integrate Big Data:** Connect to a variety of data sources with more than 800 data connectors, with native support for Hadoop, MapReduce and NoSQL. Delivers clean, consistent assets at a big data scale.

**Cons**: Isn’t resource optimized

1. Informatica

Informatica PowerCenter offers high-performance data integration to power enterprise deployments and project management. It extracts, transforms and loads data according to the organization’s requirements into data warehouses for further analysis.

**Pros**: **Perform Self-Service Analytics:** Agile processes offer users of all technical levels access to business intelligence, and role-based tools provide timely delivery of data. A codeless interface delivers a palette of pre-built graphical tools and transformations to make it easy for users to perform tasks.

Cons:

1. Lumada data integration

Part of the DataOps Suite, Lumada Data Integration deploys data pipelines at scale. Drawing large, complex assets from data lakes and warehouses, teams can orchestrate data flows across environments.

**Pros**: **Dynamic Authoring:** Update transformation jobs while the dataflow is in progress with the Data Transformation Editor. Access design tools on PDI directly from the DataOps suite (LDOS).

1. Azure data factory

Azure Data Factory is a fully managed, serverless data integration service that helps enterprises build ETL/ELT workflows. Businesses can orchestrate and automate data movement and transformation by creating end-to-end data pipelines.

**Pros**: **Debugging:** Debug the data pipeline as a whole or in parts — set breakpoints on specific workflows.

1. Infosphere Information server

Infosphere Information Server supports enterprise-level data integration and governance. With scalable information architecture, it delivers project results faster while maintaining data quality and integrity.

**Pros**: **Metadata Repository:** Share imported metadata and other assets in any server component across. the organization. Stores project configurations, reports and results for all the server’s components in one unified repository

ELT TOOLS

1. Hevo

[Hevo](https://hevodata.com/) allows you to replicate data in near real-time from [150+ sources](https://hevodata.com/integrations/pipeline/) to the destination of your choice including Snowflake, BigQuery, Redshift, Databricks, and Firebolt. Without writing a single line of code.

**Pros**: Easy config and setup with just a few clicks

**Cons:** High CPU usage

1. Luigi

Luigi is a Python library that provides a framework for building complex data pipelines. Luigi was built at Spotify. The purpose of Luigi is to allow you to automate and chain batch processes.

**Pros:** Supports dumping data to and from databases

**Cons:** Steep learning curve. You would need you to invest heavily in engineering resources that can build and maintain this infrastructure.

1. Blendo

Blendo is known as one of the best **ELT tools** enabling customers to centralize all of their different datasets and data sources into a central location. They are in the business of building new connectors and maintaining the ones already created.

**Pros:** Automated schema migrations

**Cons:** Blendo works as a extract and load kind of a set up. They do not provide a way to transform the data before or after loading to the warehouse. This becomes limiting when your ETL use cases start to evolve.

1. Matillion

Matillion is one of the best ELT Tools that is built specifically for Amazon Redshift, Azure Synapse, Google BigQuery, and Snowflake. Matillion has an ELT architecture. It sits between your raw data sources (internal, external, and third-party data) and your BI and Analytics tools.

**Pros:** Pay-as-you-go model with no long-term financial commitments.

**Cons:** You can sometimes encounter validation failure in scheduled jobs for no discernable reason.

1. StreamSets

Streamsets is a cloud-first enterprise ELT tool for extracting data from SaaS applications and databases into data warehouses and data lakes for last-mile analysis. Enterprises use Streamsets to consolidate dozens of data sources for analysis.

**Pros:** Ability to replicate, merge, as well as segment and route data.

**Cons:** It is not a good fit for very low latency use cases as it is much better suited for batch data processing.