Top 7 ETL Tools for 2021

SIGMOID

Apache Spark • Fivetran • informatica powercenter • MATILLION • sprinkle • talend • Xplenty

---

Apache Spark

Is a distributed open-source platform used for querying, processing and analyzing big data. The platform is infinitely scalable which makes it an ideal choice for even big tech giants like Apple, Microsoft and Facebook.

- The platform is way faster than other conventional large scale data processing platforms
- It provides an easy to use interface where collaborators can write Spark applications in Scala, R and Python
- It comes with pre-built libraries for Steaming, Graph and SQL computations

Fivetran

Creates a single, robust data warehouse by integrating data from SaaS services and data warehouses like Redshift, Snowflake, Azure and BigQuery. It supports a wide range of data sources, with multiple SaaS sources and allows native custom integrations.

- A fully-managed data pipeline with an intuitive web interface
- Secured data transformation and integration
- Centralized data management with no data limit

informatica powercenter

An enterprise cloud data management solution that supports end-to-end data integration lifecycle. Being database-neutral it allows seamless compatibility while working with different sources of data, including SQL and non-SQL data repositories.

- Readily available and easy training modules
- Connects various cloud applications and is hosted by AWS and MS Azure
- Highly agile and can be integrated with other tools

MATILLION

Built for cloud data integration & transformation, it uses interactive UI and SQL, and integrates with most new-age cloud databases and analytics platforms.

- Highly intuitive and agile
- Improved user experience
- Integrates with Snowflake DB, Google BigQuery, and Azure Synapse

sprinkle

A data management and analytics platform that helps automating the process of collecting data from diverse data sources and ingesting to data warehouses for seamless reporting.

- High flexibility and scalability
- Robust ML pipelines with the help of Jupyter Notebook interfaces
- Automated discovery and mapping of data types

talend

Data management through monitoring, profiling, and integrating data. The platform is compatible with both cloud and on-premise data sources and it comes embedded with many pre-built integrations.

- A free open source tool
- The interface supports drag-and-drop
- Fast and easy deployment in any cloud environment

Xplenty

A cloud-based extract, load, transform (ELT) and ETL data integration platform that simplifies the process of integrating diverse data sources. It provides BI teams with an intuitive interface for building data pipelines.

- Actionable business intelligence through centralization of data
- An effective salesforce to salesforce ETL tool
- Use of data from Rest API by supporting connectors