**Top 3 ETL Tools**

These tools will make your life easier in 2023.

In this article, we will look at the top ETL tools that are currently on the market. We will also provide a brief overview of each tool, as well as our thoughts on their future.

**AWS Glue**

**Diagram

Description automatically generated with medium confidence**

AWS Glue is a cloud service that helps you prepare and load your data for analytics. It can crawl your data sources and construct a data catalog using pre-built classifiers. It then automatically generates ETL (extract, transform, and load) scripts that you can use to move your data to the data target of your choice.

AWS Glue is a serverless data integration service that makes it easier to discover, prepare, move, and integrate data from multiple sources for analytics, machine learning (ML), and application development.

**Use-cases:**

* AWS Glue can run ETL jobs automatically as new data arrives. For example, you can configure AWS Glue to trigger ETL jobs to run as soon as new data becomes available in AWS S3 bucket.
* AWS Data Catalog to quickly discover and search multiple AWS datasets without moving the data. Once the data is cataloged, it is immediately available for search and query using Amazon Athena, Amazon EMR, and Amazon Redshift Spectrum.
* AWS Glue Studio makes it easier to visually create, run, and monitor AWS Glue ETL jobs. You can build ETL jobs that move and transform data using a drag-and-drop editor, and AWS Glue automatically generates the code.

# Azure Data Factory

# Graphical user interface, diagram, application Description automatically generated

Azure Data Factory is a cloud-based data integration service that allows you to create data-driven workflows to orchestrate and automate data movement and data transformation. With Azure Data Factory, you can create and schedule data-driven workflows (called pipelines) that can ingest data from disparate data stores, process and transform the data, and publish the results to a destination data store.

The key benefits of Azure Data Factory are:

* Highly scalable and reliable data processing
* Cost-effective data processing
* Flexible data processing
* Easy to use and manage.

# Google Cloud Dataflow

# ETL on Google Cloud with Dataprep | by Muhammad Balogun | Geek Culture | Medium

Google Cloud Dataflow is a fully managed cloud service for developing and executing a wide range of data processing patterns including ETL, batch computation, and stream processing. It is built on top of Apache Beam, an open-source framework that provides a rich set of SDKs and execution models to define data processing pipelines. Dataflow provides a simple, powerful programming model as well as the ability to automatically scale pipelines to accommodate changing data volume and processing requirements.

Some key features of Google Cloud Dataflow include:

* Fully managed data processing service
* Automated provisioning and management of processing resources
* Horizontal auto-scaling of worker resources to maximize resource utilization.
* OSS community-driven innovation with Apache Beam SDK
* Reliable and consistent exactly once processing