Forget ELT and ETL Pipelines – You Need DataOS

Modern

As enterprises modernize their data architecture, they have to make challenging decisions about legacy systems. Not all organizations are ready to offload their ETL pipelines. Unless decision-makers are also tech experts, they may not realize that ETL pipelines continue to exist within their data stack. However, digital transformation requires a comprehensive look at what's going on underneath the hood, so let's explore the humble ETL pipeline and its place in the modern data world.

What is an ETL pipeline?

ETL stands for "Extract, Transform, Load." An ETL pipeline enables data analysis behind-the-scenes and offers a repeatable process to unify data from multiple sources. It operates in three stages:

1. **Extract**: Data from all involved sources moves from the source to a staging area. Since it may not be possible to move all data from all sources at the same time, the staging area helps reduce bottlenecks.
2. **Transform**: Raw data becomes data models—structured, cleaned, and formatted without duplicates or incorrect entries.
3. **Load**: Now that the transformed data is in a usable format, it can be uploaded to the warehouse.

ETL pipelines have been in use since the 1970s and have changed a lot in the way businesses derive value from data. If an organization has been around for a few decades, there's a good chance ETL pipelines are present in the data stack.

ETLs can cause trouble for organizations in digital transformation

Originally, ETL pipelines changed data analysis because they automated much of the mundane work required to retrieve and scrub data. Most data fell into a structured model originating in databases and operational systems designed for data like this. ETL pipelines were formatted for specific users and allowed IT to avoid tedious code-writing just to query data.

Data today isn't so structured. Organizations are harvesting data from social media, images, voice recordings...nothing is "neat" about data anymore. It's bigger and messier and only becoming more so as people live out their lives online.

ETL pipelines can't keep up. They're too rigid and don't transfer well from one set of users to another or from one type of data to another. They require a heavy investment from IT and slow down insights needed for decisions today, not months from now.

Some organizations replaced ETL with ELT, where data loading and transforming are decoupled to reduce those bottlenecks. However, even ELTs — while more efficient than ETL pipelines — still have one key obstacle for a truly transformed data stack: they don't facilitate self-service for the nontechnical user.

Enterprises can enable data self-service with a data operating system

ETL and ELT pipelines require technical know-how to build, maintain, and troubleshoot, but companies need everyone in the organization to become data stakeholders. Departments don't have time to wait months to receive data-backed insights to make decisions. They need access to data now, and they need the support to build reports and queries without overloading IT with requests.

A data operating system like DataOS from The Modern Data Company (Modern) offers a new way to manage data. It's a connective tissue for an entire data ecosystem, uniting tools and data sources from all over the enterprise. Nontechnical users can search for the data they need using a simple, Google-like approach and work directly with data without making copies or moving it.

Without data movement, enterprises experience fewer risks in security and can implement organization-wide governance policies that remain consistent no matter the user. For example, the marketing department can view critical marketing data with columns of sensitive information removed through a marketing Data Lens.

Legal departments can see all columns and rows through their own Data Lens.

Both departments can accomplish their own data transformation through a right-to-left engineering approach that abstracts data complexity — the system removes the need for specialized technical knowledge to build the right pipelines. Users drag and drop the functions they need, and DataOS takes over to create exactly the right thing.

Move from pipeline usage to data ownership

ETL pipelines were a significant improvement over hand-coding, but DataOS takes data analysis to the edge with a comprehensive, connective, and consistent data solution. Business users gain insights in near real-time because they're taking ownership of the data process themselves. Organizations can finally free data for all stakeholders safely and efficiently, and IT no longer needs to build tedious ETL or ELT pipelines.

To see it work specifically for your organization and to understand the scope of what DataOS can do for your organization's data, download our latest white paper: "A Modern Data Strategy for Enterprises".

See more →

BY E. WALLACE