

More useful data to predict customer behavior and deliver better experiences

How the data team at Autodesk simplified data preparation and analytics for the enterprise with Matillion ETL and Snowflake

If there's one thing most technologists can agree upon, it is that technology for technology's sake doesn't always live up to the promise of a better life. The data team at Autodesk firmly believes that adopting big data trends such as data lakes and data warehouses should not only improve their end user's situation, but also benefit themselves. Achieving this requires technologies and tools that are simple to use, yet powerful and effective in getting the job done right, thus improving the lives of everyone involved, for the better.

About Autodesk

Autodesk makes software that helps customers to design and make a better world for all. Its technology spans architecture, engineering, construction, product design, manufacturing, media and entertainment, empowering innovators everywhere to solve challenges big and small.

Industry:	Computer Software
Company size:	10-15,000 employees
Headquarters:	San Rafael, CA
Type:	Public Company Founded 1982

Providing self-serve data insight for a range of users

Like most modern data-savvy organizations, Autodesk is on a journey to fully exploit the data that the business generates, converting it into new opportunities for growth and long-term sustainability. This means that the data team is under continuous pressure to transform data faster, make it available more quickly, and build solutions that operate in near real time.

The Autodesk Data Platform is the largest data lake at Autodesk, used by people across the enterprise in a self-serve manner to obtain insight into product usage from instrumentation data being captured and stored on the cloud. Users from across the

business can perform analytics on shared data sets, or can combine their own data to satisfy advanced requirements. The Autodesk data team needs to support a diverse group of consumers who need the ability to build something unique to their own needs, but on a common platform.

Without this important instrumentation data, users wouldn't be able to produce accurate scoring models, negatively impacting the ability to generate quality new leads. Says Mark Kidwell, Chief Data Architect, Autodesk Data Platform: "Our Salesforce salespeople would just be sitting there waiting for more stuff to do, people to call."

As with most business critical initiatives, speed is key in the data transformation arena. Autodesk needed to improve both the velocity and quality of data insights to the business. One of the data team's biggest challenges was the huge range of skill levels among their own staff and their end users. To deliver on their strategic data goals, they needed a solution that could:

- ✓ Fully exploit business data to drive growth and long term sustainability
- ✓ Provide self serve data from a common platform that is up to date and high quality
- ✓ Cater to user needs with diverse skill levels across the extended organization

A low-code/ no-code solution that leverages the power of Snowflake

After evaluating several technologies against their use cases, Autodesk chose Snowflake as their common enterprise data platform of choice. The data team then embarked on a data integration project that took nearly a year to implement. Because Snowflake was an integral part of their data engine and tool stack, Autodesk needed an ETL solution that would facilitate both rapid implementation and frictionless adoption, making the most of the native power and features of Snowflake.

Early on, Autodesk identified a need to simultaneously support two approaches to data pipeline automation: a "low-code" solution to cater to their more advanced needs as well as a "no-code" managed service. Autodesk couldn't compromise on some of the sophistication offered by custom coded solutions, even though a key criteria for a new solution was ease of use, to facilitate rapid adoption and faster time to value. This dual challenge quickly eliminated ETL providers such as Informatica, which was based on older ETL technology and too hard to use for all but the most technical users. In fact, many ETL tools Autodesk

considered relied too much on users acquiring (and maintaining) scripting skills. Matillion ETL's visual design approach clearly offered a better alternative to a wider range of users. Today, Autodesk's data integration team of about 10 people uses Matillion ETL for Snowflake extensively, citing the ease of use and flexibility to support both ELT and ETL paradigms as some of the key reasons why.

The Autodesk data team historically used a high proportion of highly technical staff to do complex coding work on the data lake. However, the company soon realized that it could not sustain that level of staffing and maintenance. The team needed to take a no-code approach for the bulk of data ingestion, combined with a low-code approach for more complex data transformation work.

“Just really simple things were really hard for us and bringing in the right engine and the right tools has helped fix that.”

– Mark Kidwell, Autodesk

Matillion ETL combines both ease of use and support for complex transformations, by hiding the complexity unless required by the user. This elegant approach is a key enabler in overcoming the dual challenge of simultaneously providing ease of use and sophistication. The switch to Matillion ETL also enabled Autodesk to attract and quickly onboard new talent more familiar with modern visual design paradigms such as drop-down boxes and drag-and-drop features.

Features used by Autodesk

- ✓ **Visual design**
Simplifies the process of creating data ingestion and transformation business rules, making it accessible to a wider range of user skillsets and easier to maintain
- ✓ **DB Query component**
Enables access to an extensive range of databases and APIs
- ✓ **Dynamic Variables and Bridge components**
Makes jobs reusable, reduces maintenance overhead, simplifies troubleshooting, and standardizes operations
- ✓ **Unloading data from Snowflake**
Populates the data lake with data transformations done in Snowflake, using Matillion as a “wrapper”

Providing self-serve data insight for a range of users

Using Snowflake and Matillion ETL, Autodesk has been able to achieve two important business data goals.

First of all, the new solution has enabled the data team to provide transformed, analytics-ready data that data scientists, such as financial analysts using the financial reporting platform and data mart, can use in models to more accurately and effectively score leads and predict churn and customer losses. Matillion is used to bring the data in, transform it, unload it, and make it ready all the way through.

Second, Autodesk has vastly improved the velocity and quality of enterprise data by retiring legacy technology that was responsible for 90 percent of their data quality issues. The previous experience, built around Spark and legacy ingestion tooling, was difficult to use and error prone. By adopting Matillion ETL, Autodesk simplified their entire process of ingestion, transformation, and analytics.

Looking forward: Data capabilities for everyone

Autodesk is continuing to expand its adoption of Snowflake to allow more users than ever access to useful data. The company is now exploring the possibility of making Matillion available as a platform-level offering for any staff member who wants to adjust their own data, in the same way the data team and data scientists do today. By increasing the availability of data and analytics across the business, Autodesk hopes to increase the value of timely, quality enterprise data availability across the organization.

About Matillion

Matillion ETL is a cloud-native data integration platform that helps modern teams transform their business with data. With a sophisticated yet easy to use visual approach, Matillion ETL provides a low-code interface to turbocharge cloud data ingestion and transformation workflows, and give data professionals the cost effective tools they need to get results faster in the cloud. As a truly cloud-native solution, Matillion ETL is the only platform designed to leverage all of the unique capabilities and services in your cloud data platform of choice. And with enterprise grade security and collaboration features included at its core, Matillion ETL helps a new generation of workers unleash the untapped potential of their data. **Learn more at matillion.com**