

Automated Lineage Visualisation: Global Retailer

CHALLENGE

When a Global Retailer spent several billion dollars to take over an e-commerce start up in 2016, the financial media was quick to raise their eyebrows.

Since completing the deal, the retailer's e-commerce sales have soared, climbing 63% in its most recent quarter. Online inventory has also grown, from just 10 million items last year to at least 67 million today.

However, with growth came complexity. Sources of data became unclear, bringing difficulty in understanding how to aggregate the data and determine from which system(s) the information originated.

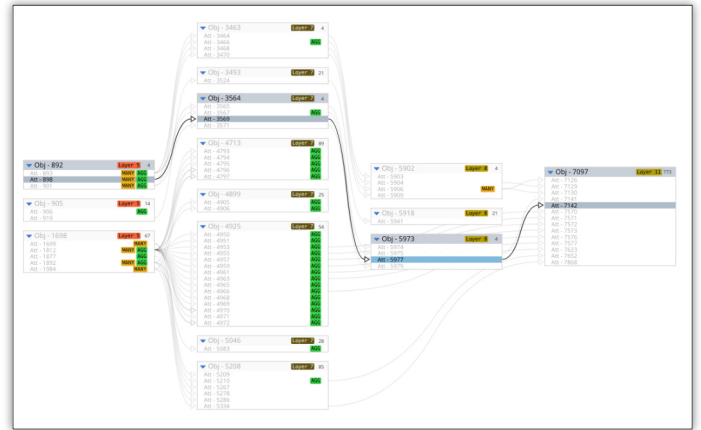
In early 2018, the acquired e-commerce start-up arm of the retailer initiated a search for a solution that could provide a visual, searchable and transparent view of the input/output pipelines from their Data Platform. A solution that would enable them to trace data back to its origin, obtain impact assessments of modifications to upstream systems and provide greater quality assurance. They were very clear from the outset that they wanted to build lineage automatically from source, with unlimited potential to scale.

SOLUTION

After an exhaustive search of the market, Solidatus was selected. A Senior Director at the retailer said, "We have responsibility for production data across all systems. Our system is the common platform that publishes and consolidates into a central location, building up data-marts in the key business domains". He went on to say, "We did our due diligence and there are very few products out there that do lineage well".

The retailer is using Solidatus to fully automate the storage, versioning and visualisation of their field-level data lineage.

When their jobs build-up, they expose descriptors about the inputs and outputs of the job, which are then posted to the Solidatus API in the cloud. Solidatus compares, maintains and versions these data flows, allowing them to query and visualise their end-to-end data flows.



RESULT

Solidatus is being widely used to support and connect key business domains, providing an interface with analytical layers. "It provides a clear visualisation of our dependencies and reflects the current build status with the objective to eliminate end-to-end failures. It has fully met our primary use case of using Solidatus to visualise the current production environment and there is now scope for more usage as part of our change management planning process and in many other areas", comments a Senior Director at the company.

KEY POINTS

- Industry standard and secure RESTful HTTP API.
- Automate data lineage discovery, storage and retrieval.
- Export data from Solidatus to generate reports.
- Integrate Solidatus into your existing governance ecosystem.
- Bootstrap Solidatus by importing existing metadata.

ABOUT SOLIDATUS

Solidatus is a specialised, powerful and modern data management tool. The simple, intuitive and flexible web-based application allows organisations to rapidly discover, visualise and understand how data flows through their systems and the relationships it has.