



# Utilities – Metadata-Driven ETL Solution



# Utilities – Metadata-driven ETL solution

## Automated ETL through metadata-driven framework using Azure technologies

### PROBLEM

A utilities provider approached exposé to assist as they transitioned out of an on-premise call centre system (Genesys) and into a new cloud-based system (Genesys Cloud). The on-premises system was used as the primary data source for several reports that the business utilises. Extraction and transformation of the data were manual and laborious and required numerous hours to complete.

Exposé was tasked to first understand how data can be extracted out of the new system through API endpoints. Exposé then needed to build an automated ETL process to take the data out of Genesys Cloud and store the data in a Data Lake. The solution being built also needed to adhere to the architecture of the client's existing Data Platform.

Once the data had gone entirely through the ETL process and authoritative entities had been established, Power BI data models needed to be created to enable the Power BI reports designed by the internal team.

### SOLUTION

Exposé worked closely with the client's internal team to identify various API endpoints needed to build authoritative entities used in the data models. APIs were heavily analysed to explore their limitations, parameters, dependencies and resulting data.

Exposé used a metadata-driven framework to orchestrate the ETL process. This approach provided flexibility to the client as new data sources were identified and added or removed if necessary.

Exposé developed bespoke and patterned pipelines in Synapse Analytics to orchestrate data extraction out of the source system and storage in the client's Data Warehouse.

Databricks were used for incremental loading, flattening of the extracted JSON files and business transformations to create the authoritative entities used in the data model.

Transformed data was then loaded into tables in Synapse Dedicated SQL Pools, where the Power BI Datasets and Dataflows consumed it as data sources in Power BI Reports.

### BUSINESS BENEFITS

The solution provided the client with an automated ETL process that has the flexibility to add more data sources or remove existing sources if they wish to. The solution was tailored to fit their existing data platform, so no rework was needed, and existing technologies were leveraged, making the solution easy to maintain.

The reporting team also felt tangible benefits:

- The reporting team are now focused on analysing data and drawing insights rather than creating reports.
- Data is always readily available for anyone to access and always up-to-date.
- The solution prompted momentum and significant interest in expanding to other datasets in Genesys Cloud to be extracted and analysed so that the reporting team can draw insights on areas the business never had visibility on.