kyndryl

Unifying enterprise-wide data to unlock innovation through analytics

Leading multinational energy sector group | Energy

Business opportunity

An energy group present in more than 30 countries across four continents is seeking to accelerate its diversification from gas and electricity into renewables. At the same time, its customers are changing how they consume and pay for energy, while becoming increasingly keen to minimize their carbon footprint.

To complete the transformation will require innovation, and the group will need to make better use of data analytics to drive informed decision-making and strategic planning. By creating a holistic view of global data, the group would gain the insight it needed to ensure a profitable yet sustainable future.

Technical challenge

With data locked away in multiple on-premises systems around the world, the group struggled to create complete and consistent sets of data for analysis. Poor data quality made it impossible to run integrated analytics, and the company could not easily identify trends, patterns or opportunities to drive growth.

In operational terms, the lack of a centralized data repository meant that IT and business personnel were spending large amounts of time finding, managing and preparing data for analysis. These inefficient data-management processes raised questions about the consistency of datasets, making it difficult to generate accurate, reliable or meaningful conclusions

Our solution

Together, Kyndryl and the energy group built a cloud-based repository of data for analytics and machine learning.

Data is sent via SFTP from on-premises silos to an AWS data lake. Its arrival on Amazon S3 triggers an Apache Airflow workflow that segments and normalizes the data to ensure consistency and reliability. This unified dataset is stored in Apache Parquet format, ready for analytical processing by Amazon EMR (Amazon Elastic MapReduce).

For machine learning — which the energy group uses to identify high-value customers — Kyndryl implemented an XGBoost model on Amazon Sagemaker, which uses the EMR cluster as its data source.

The power of partnership

Together, Kyndryl and the energy group built a cloud-based repository of data for analytics and machine learning.

Data is sent via SFTP from on-premises silos to an AWS data lake. Its arrival on Amazon S3 triggers an Apache Airflow workflow that segments and normalizes the data to ensure consistency and reliability. This unified dataset is stored in Apache Parquet format, ready for analytical processing by Amazon EMR (Amazon Elastic MapReduce).

For machine learning — which the energy group uses to identify high-value customers — Kyndryl implemented an XGBoost model on Amazon Sagemaker, which uses the EMR cluster as its data source

What progress looks like

With a single, centralized repository for data, the group can execute sophisticated analytics and modeling rapidly and easily, generating consistent results. Data-driven insight helps the group innovate towards a more sustainable and profitable future by enabling it to identify obstacles and opportunities.

→ Cut operational costs by 40% through centralization

→ Improved quality and consistency of data

→ Reduced manual effort in finding and querying data

→ Enabled faster, higher-quality decision-making

Beyond the initial pilot, Kyndryl is now advising the retailer to optimize date use and explore IOT opportunities for food safety. Opportunities include smart devices prompting cashiers to refill coffee machines and re-stock refrigerators.

What's your next
digital business challenge?

Let's tackle it together. →

kyndryl.