

AWS-based end-to-end automation and reporting

Client

An American biopharmaceutical company headquartered in California with offices in more than 30 countries

Company Size

17,000+ employees

Location

Global

Featured Partners



The Swiss division of the global pharmaceutical company sought to automate the data collection and reporting process.

The company partnered with our data strategy team to develop a fully automated solution in AWS using Lambda functions, a PostgreSQL database hosted on AWS RDS and Python. Reporting is now fully carried out in Power BI, taking advantage of the PBI Rest API to ensure seamless data refreshes. Data is automatically collected from SFTP, Starburst, emails, and S3. It is then processed and stored in bronze, silver, and gold layers. The project contributed to a **50% lower licensing cost.**



Process Automation



BI & Visualizations



Advanced Analytics



Data Cleansing & Quality

The Challenge: Overcoming inefficient data collection and reporting



Objective: Automate and streamline the client's reporting process to reduce manual workload, improve accuracy, and increase data refresh frequency.



Existing Issues: Reports relied on manual data downloads from multiple sources, making the process time-consuming, error-prone, and limited to monthly updates.



Solution Needed: A fully automated, scalable, and flexible reporting pipeline capable of integrating multiple data sources and supporting more frequent refreshes.



Outcome: The new AWS-based solution enabled daily reporting, reduced costs, and improved data accuracy.



The Swiss division of a pharmaceutical company relied on manual, error-prone reporting processes that limited insights to a monthly basis and consumed significant time and resources. By implementing an automated AWS-based solution with PostgreSQL, Python, and Power BI, the company achieved daily reporting, improved accuracy, and significantly reduced operational costs.

The Solution: AWS-based end-to-end reporting

We developed a fully automated reporting solution in AWS, which enabled faster and more accurate insights by eliminating manual data extraction. The platform integrated multiple data sources into a scalable cloud environment and stored data in bronze, silver, and gold layers, delivering daily reports through Power BI for improved consistency and efficiency.

Phase 1: Data Collection and Ingestion

The team automated the extraction of data from multiple sources, including Salesforce, IQVIA, SFTP, emails, and S3, ensuring reliable and scalable ingestion.

Workstreams:

- Data ingestion
- Source connectivity setup
- Initial data validation

Phase 2: Data Processing and Storage

Data was processed and organized into bronze, silver, and gold layers within a PostgreSQL database on AWS RDS, enabling structured and consistent storage.

Workstreams:

- Data cleaning and transformation
- Layered storage implementation (bronze, silver, gold)
- Database setup and optimization

Phase 3: Reporting and Automation

Automated reporting was implemented in Power BI using the REST API, providing daily refreshed reports with improved accuracy and consistency.

Workstreams:

- Power BI report development
- API-based data refresh automation
- Reporting validation and monitoring

Services and Technologies Used:

Services:

- Process Automation
- BI & Visualizations
- Advanced Analytics
- Data Cleansing & Quality

Technologies:

- Python
- SFTP
- Starburst
- AWS Lambda, RDS, S3
- Microsoft Power BI
- Power BI REST API
- Power Automate

The Results: Impact on the client organization

The project led to a more efficient reporting process, significantly reduced licensing costs by around 50%, and improved the company's ability to access timely and accurate data. By automating data collection and processing, the company achieved daily insights, enhanced consistency across reports, and streamlined overall analytics workflows.



Daily Reporting Enabled: We increased reporting frequency from monthly to daily, providing the client with timely and actionable insights.



Cost Efficiency Achieved: This project reduced licensing and operational costs by 50% through streamlined processes and automation.



Manual Workload Eliminated: Automated data extraction and processing significantly reduced the time spent on repetitive manual tasks.



Scalable Cloud Solution: Implementing AWS services allowed easy integration of new data sources and supported growing data volumes.



Data Accuracy Improved: By centralizing and automating data handling, we drastically minimized reporting errors.