

Serverless Data Lake for Life Sciences

Accelerate Biomedical Insights with Low-Code Data Solutions

The Challenge of Biomedical Data Utilization

The life sciences industry is generating biomedical data at an unprecedented scale, yet many organizations struggle to fully leverage this data. With data growing exponentially, efficient management and analysis become critical to maintaining a competitive edge in drug discovery and development. Without a streamlined data architecture, life sciences companies risk slower time-to-insight and reduced innovation capacity.

Enhancing Clinical Data Analysis

Our Catalyst provides a tailored, no-code solution designed to manage and analyze vast biomedical datasets efficiently. By leveraging the AWS suite of data storage and analytics tools, we help life science companies rapidly transform their raw data into actionable intelligence for a decisive edge in the competitive landscape. Our solution integrates diverse data sources, enabling teams to seamlessly scale their data infrastructure while empowering scientists to wrangle and explore data without the need for complex coding.

Key Activities

Discovery & Planning

Conduct a workshop to define your data lake foundation and data exploration needs, capturing key requirements for data import and access patterns

Design & Implementation

Design the infrastructure, data ingestion, and security model based on your needs, then build the data lake with up to three schema sources

Enablement

Run an exploratory data analysis (EDA) lab using AWS Glue DataBrew and Lake Formation to equip your team with hands-on experience in data exploration, cataloging, and governance

Highlights

- **Accelerated Time to Market:** Leverage AWS serverless services like Lake Formation, S3, and Glue to build low-code data lakes that significantly reduce time to market
- **No-Code Data Analysis:** Run seamless data analysis without coding, using Amazon Athena and AWS Glue DataBrew to simplify exploration and insights
- **Secure, Scalable Infrastructure:** Easily scale data sources and pipelines with prebuilt Infrastructure as Code (IaC) automation, while ensuring data security and governance through AWS Lake Formation

Deliverables

- **Workshop & Documentation:** Participate in a workshop to explore your data landscape and receive comprehensive documentation of requirements
- **MVP & Architecture Designs:** Receive an MVP data lake with detailed architecture diagrams, data workflows, and technical documentation
- **EDA Enablement Lab:** Gain hands-on experience with an EDA lab, focusing on data wrangling, stewardship, and governance, supported by custom templates and scripts