

Breaking Barriers in the use of Biomedical Data- Multi-modal Data Management

Agenda

- Introducing Polly: Platform for managing multi-modal biomedical data
- Example of how to use Polly: Analyzing Metabolomics Data
- Discussion



*"The value is in the data, it is not in the tools. That is the one thing, it's a bit of a hobby horse for me. One thing I always point to in these discussions around data, don't underestimate the amount of time and value in doing what is really often difficult and not so rewarding directly work, like **cleaning data sets isn't always fun, but it is often the most valuable thing you can do.**"*

-Dr. Jeffrey Reid,

Regeneron's Chief Data Officer

Leaders in Multi-Modal Biomedical Data Management

with Differentiated Technology & Expertise



#1

30+ customers deriving value from our technology and expertise.

#2

A malleable cloud friendly technology that can seamlessly fit into your R&D Ecosystem.

#3

With a team that can quickly assimilate industry domain into high value services.

2023 Frost & Sullivan Enabling Technology Leadership Award

The North America Tech-Enabled Drug Discovery Industry

Industry Research Analysis / 9B43 / B8

Published: 2023-09-19



**Most
Innovative
Companies**

FASTCOMPANY

Polly, A Scalable Platform for Multi-Modal Data Management



Data Harmonization & Ingestion



Polly Pipelines

Workflow orchestration system(supports Nextflow)



Polly Curation

Manual and automated data curation



Polly Sync

Import files from many external sources

Data Storage & Management



Polly Atlas

For storing structured data in a data model



Polly Workspaces

For storing unstructured data

Data Consumption



Polly Apps

Custom apps running on a highly scalable infra (Hail, IGV)



Polly Notebooks

Data exploration/analysis using Jupyter notebooks

25+

Data types Supported

PBs

of Data stored

10+

Integrations to ML, BI & Analysis Tools

35 TBs

of datasets processed and harmonized / month

< 15
secs

To query millions of records at once

17+

Proprietary Web-Applications developed

Polly brings value on day 1 and can evolve as your analytics needs evolve 

#1

**Managed
Cloud
Platform**

#2

**In-built Data
Security &
Governance**

#3

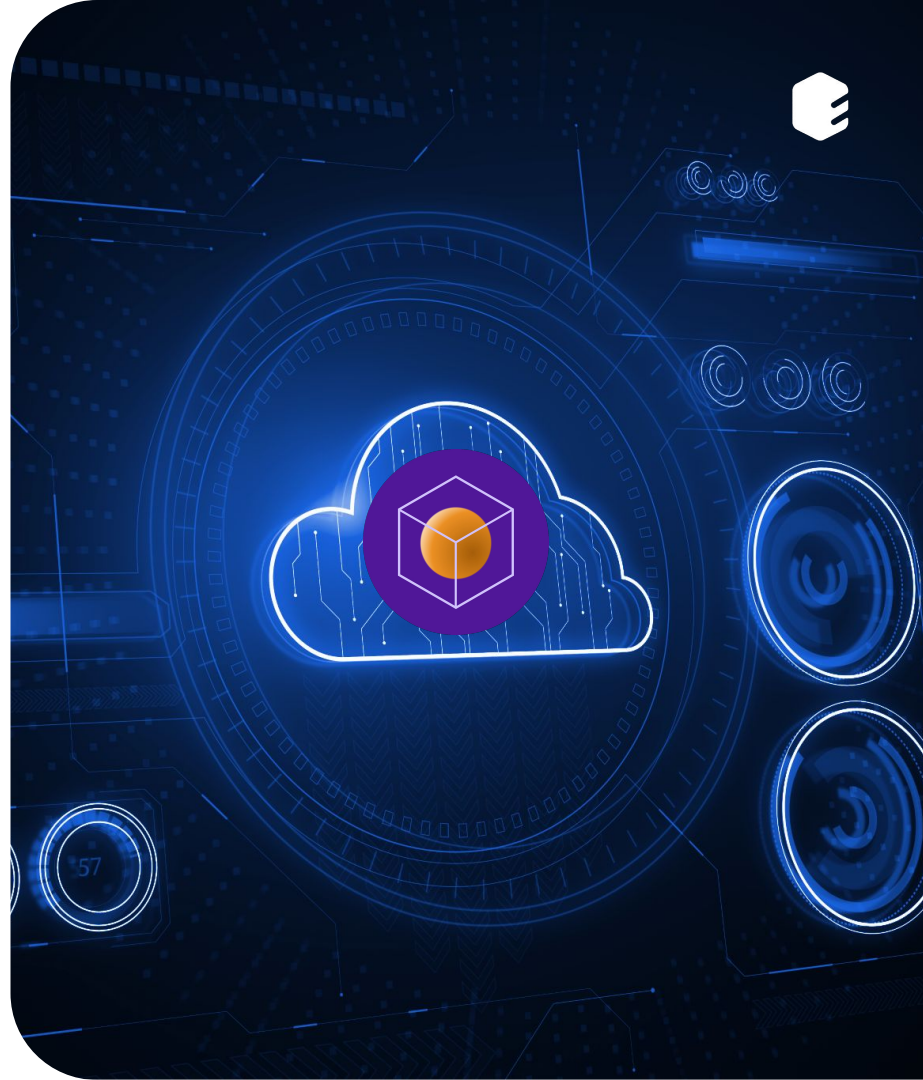
**Raw multi-omics
Data to Insights**

With Polly - Fully Managed Cloud Platform From Day 1

*Building this in-house could take **2 years, 15+ Engineers** and **millions of dollars***

Benefits of a fully managed cloud platform:

- New institutes can be **easily onboarded** on the same platform
- Support and maintenance included, with **99% uptime.**
- Secure collaboration across users, multiple sites and stakeholders with **granular access controls**



With Polly - **In-built** Data Security & Governance

*Baking in security & compliances into a data platform takes an **additional 6 months***

With Polly your data is secure from day 1:

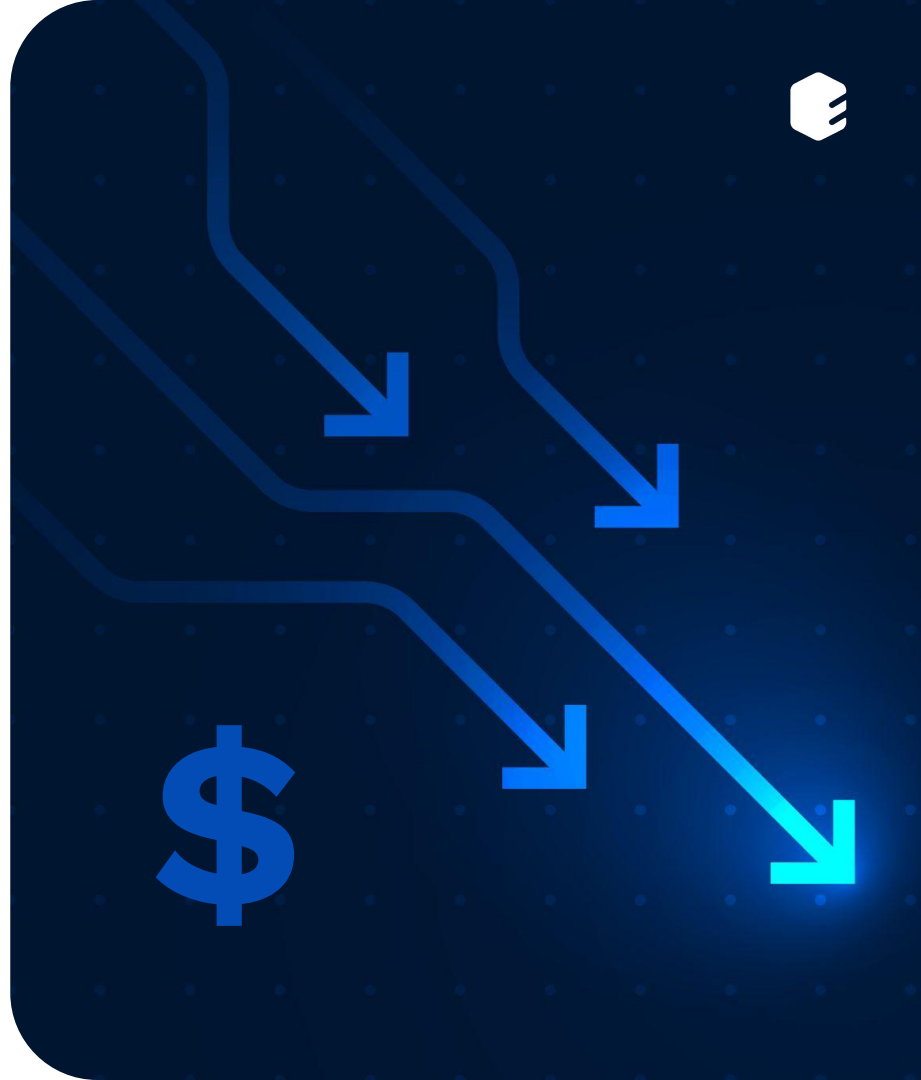
- Compliances Supported: **SOC 2, GDPR, HIPAA**
- Platform supports **single sign on (SSO)** services for multiple simultaneous users and role based access control.
- Enterprise-grade Authentication and Authorization.
- Deployment **not dependent** on geography.



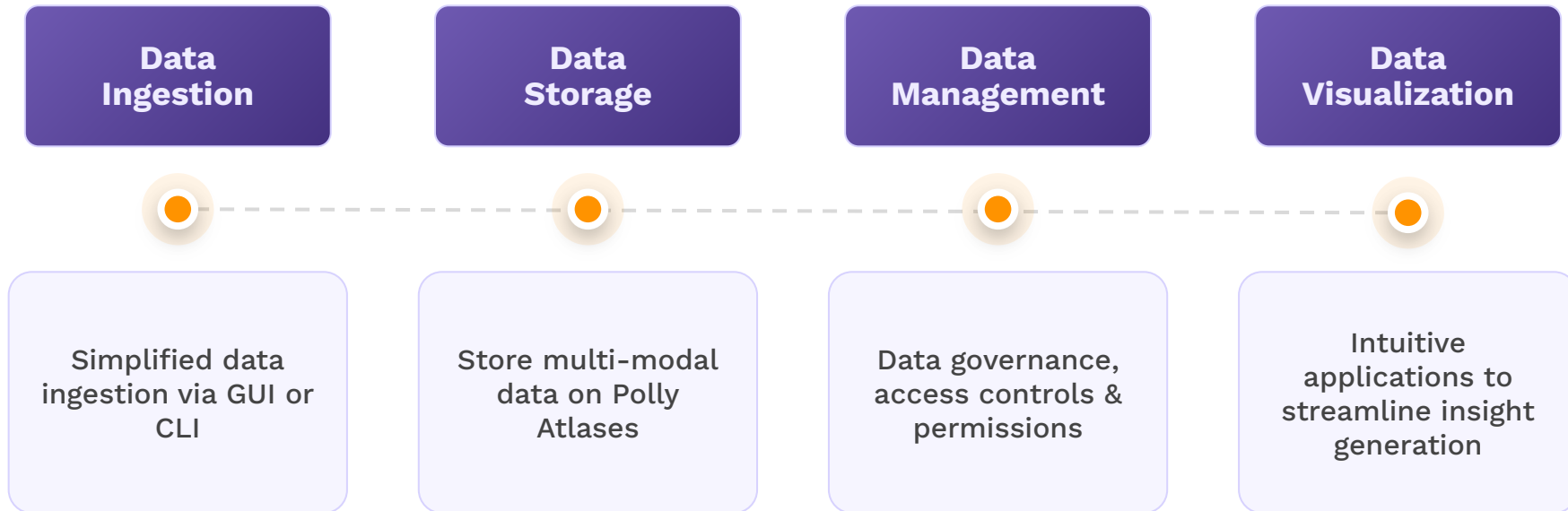
With Polly - Accelerate Raw Data to Insights by Multi-Fold

Scalable Infrastructure that supports:

- Harmonization & Analysis of **25+ Data Types**
- **2X Cheaper data processing** than vanilla AWS
- Ingestion, processing & harmonization of **5000 samples per week**
- Seamless Integration of **new applications or pipelines** on Polly's cloud infrastructure.



From Raw Data to Insights: Data Journey on Polly



Eliminate Manual Data Transfers with Polly Ingestion



Data
Ingestion

Data
Storage

Data
Management

Data
Visualization



Custom GUI for Data Ingestion

Overview

Upload Data

Download Data

Save The State

File Upload

Upload multiple files of different data types to your Dashboard. Fill the form below to add a file. To add multiple files click on 'Save & Add another file' button and fill the form again. Once all files have been uploaded, click on the 'Upload' button to finish the upload process and add your files securely to your Polly Storage.

Select Datatype

- Clinical Metadata
- Clinical Metadata
- Elispot
- ICS
- IHC-Polaris
- Cytokine
- ctDNA
- SOD

Upload ORRECIST

Additional Comments

Save & Add Another Datatype

If you wish to notify other users when the files are uploaded, please select the 'Notify additional users' options and add the email addresses of multiple user. (separated by commas)

Notify additional users (Default User: tech_sales@elucidata.io)

Upload

5000 samples can be ingested weekly

Pipeline Monitoring

Run Analysis

All Runs (1275)

Filters

Run ID	Created by	Status	Batch	Pipeline	Start Date & Time	Time Elapsed
XCKLD10124	JD Jainik Dedhia	Pending	Bulk RNA Seq 22 02	Bulk RNA Sequence v8	22/02/2024, 11:06 AM	01:29:31 Hours
PODRT18124	MS Manimala Sen	Completed	Bulk RNA Seq 22 02	Kalisto v1.11	22/02/2024, 11:06 AM	01:29:31 Hours
XCKLD10124	AJ Abhishek Jha	Completed	Bulk RNA Seq 22 02	Kalisto v1.11	22/02/2024, 11:06 AM	01:29:31 Hours
ZYLLD19124	SP Swetabh Pathak	Completed	Bulk RNA Seq 22 02	Bulk RNA Sequence v8	22/02/2024, 11:06 AM	01:29:31 Hours
WEDJH09124	SP Swetabh Pathak	Completed	Bulk RNA Seq 22 02	Carcinoma Testing v12	22/02/2024, 11:06 AM	01:29:31 Hours
LOKNH07224	JD Jainik Dedhia	Completed	Bulk RNA Seq 22 02	Bulk RNA Sequence v8	22/02/2024, 11:06 AM	01:29:31 Hours
XCKLD10124	JD Jainik Dedhia	Completed	Bulk RNA Seq 22 02	Bulk RNA Sequence v8	22/02/2024, 11:06 AM	01:29:31 Hours
ULUJT18224	MS Manimala Sen	Completed	Bulk RNA Seq 22 02	AlphaFold Test v4.71	22/02/2024, 11:06 AM	01:29:31 Hours
HFSLD09224	YL Yogesh Lakhota	Completed	Bulk RNA Seq 22 02	Kalisto v1.11	22/02/2024, 11:06 AM	01:29:31 Hours
UYTLD19024	MS Manimala Sen	Error	Bulk RNA Seq 22 02	Bulk RNA Sequence v8	22/02/2024, 11:06 AM	01:29:31 Hours
PODRT18124	SP Swetabh Pathak	Completed	Bulk RNA Seq 22 02	Bulk RNA Sequence v8	22/02/2024, 11:06 AM	01:29:31 Hours

25 Rows per page

Showing 1 - 25 of 345

Harmonize & Index Data on a **Polly Atlas**



Data
Ingestion

**Data
Storage**

Data
Management

Data
Visualization



Multi-modal data is harmonized, integrated & Indexed on Polly Atlases

- Processed, annotated with critical metadata & QC-ed with **harmonization engine**
 - LLM-powered harmonization accelerates metadata annotation by **10X**
 - Values & Metadata **validated by experts**
- Common Data model provides **relational link** between Patient metadata and molecular data

Manage Sensitive Data & Projects over **Workspaces**



Data
Ingestion

Data
Storage

**Data
Management**

Data
Visualization



**Eliminate Data Silos, manage permissions
& control access to sensitive data assets**

- Unified view of **raw data, processed data & corresponding results**, in the same location
- Version control data, metadata & analyses.
- **Control permissions** & provide the **right access controls** to data & projects on Workspaces

Streamline insight generation with **Intuitive Applications**



Data
Ingestion

Data
Storage

Data
Management

**Data
Visualization**



- **Simplified data exploration** via custom filtering options on an Atlas UI
- Seamless comparison across patient cohorts with **integrated multi-source data**
- Visualize & analyze on native apps like **Phantus, IGV or Intomix**
- Connect with external tools like **Spotfire with APIs**

| **Thank You**