code ocean

CASE STUDY

How CytoReason scaled computational multiomics research while keeping costs under control

The challenge

Founded in 2016, CytoReason is a leader in developing cutting-edge computational disease models. This includes:

- Using a proprietary data and machine learning model to reconstruct cellular information from bulk tissue
- Training an immune-specific neural language processing engine
- Integrating multi-omics data from collaborators and public sources into a disease model data structure

Initially a team of five, they developed their own internal tools for managing research. When the team grew to over thirty researchers, they quickly found that their tools didn't scale and lacked critical functionality:

- **DevOps complexity:** to compensate for managing multiple cloud-based machines, they compromised on a single machine allocated to multiple projects per researcher
- **Research delays:** under the higher project load, their internal framework for development operations and preserving projects was more difficult to maintain, slowing down research and making results and data hard to find
- **Limited IT bandwidth:** Computational Biologists and algorithm developers were often frustrated, and internal IT and DevOps teams did not have the bandwidth to support

"Code Ocean opens the door to automating and scaling major aspects of the research process. With the Code Ocean platform, our scientists can focus on their core competencies, do more, think bigger, and achieve higher goals. ... Code Ocean also makes it possible to reproduce projects so we can quickly identify and address quality issues."

Yuval Kalugny
CytoReason VP Engineering

Key features used

- Lineage graph
- Capsules
- Data
- Admin panel

code ocean

The solution

CytoReason implemented Code Ocean by installing it on their AWS VPC instance:

1.

To optimize compute resources and costs, CytoReason now has multiple cloud machines allocated to projects on an on-demand basis, and scaled back by the platform when not in use

2.

Compute Capsules have created a common computational standard for work while also supporting RStudio, Jupyter, and other popular tools

3.

All computational research is now available for straightforward sharing and collaboration thanks to integrated role-based access control, as well as guaranteed reproducibility and traceability

About CytoReason

CytoReason is a leading technology company developing computational disease models. The company collects proprietary data from pharmaceutical companies and uses it to simulate human diseases – tissue by tissue and cell by cell. With CytoReason's massive database and AI-led platform, pharma and biotech companies can identify new opportunities, shorten trial phases, reduce development costs, and increase the likelihood of drug approval. To date, five of the world's top ten pharma companies use CytoReason's technology.

The results

Productivity
15-30% increase in researcher productivity

Compute efficiency
Through use of built-in compute management features

DevOps cost reduction
$100,000 annual savings realized for DevOps work

Key integrations

Jupyter

AWS EC2

R Studio

Docker

Git