CytoReason

Boost computational productivity
by streamlining research workflow and platform
for guaranteed reproducibility

CytoReason is a global leader in machine learning for drug discovery and development. CytoReason technology uses a proprietary data and machine learning model to reconstruct cellular information from bulk tissue, and train an immune-specific neural language processing engine. The company's platform integrates multi-omics data (gene, protein, cell, and microbiome) from collaborators and public sources into a disease model data structure that intuitively reflects human biology. CytoReason works with pharma clients like Pfizer and GlaxoSmithKline to deliver molecular-level insights that predict drug effects for novel targets, indications, combinations, and biomarkers.

The Challenge

CytoReason was founded in 2016 with a team of five scientists dealing with the most innovative research challenges faced by biopharma. They had technical experience to develop their own initial tools for computational research management. As their team grew to over 30, they found that these tools didn't scale and lacked critical functionality. Due to the complexity of managing multiple machines in the cloud and the risk of idle machines raising compute costs, they compromised on a single cloud machine allocated to multiple projects, per researcher. 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. Computational biologists and algorithm developers were often frustrated, and internal IT and DevOps teams did not have the bandwidth to manage issues manually.

Cytoreason wanted to focus on research, not operations. Therefore searched for an integrated solution for both computational reproducibility and automated IT operations.

Goal

Achieve comprehensive high-performance research environment to multiply productivity and guarantee data and code preservation

Key benefits

Increased researcher productivity by 15–30%

Guaranteed reproducibility

Optimized compute and DevOps costs

> 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.

Yuval Kalugny, CytoReason VP Engineering

The Solution

Code Ocean is a cloud-based platform for computational research that streamlines workflows in creating, organizing, and sharing projects, automating computing resources, and guarantees reproducibility. The platform continuously captures the exact version of code, data, and computing environment that generated every result.

CytoReason chose to adopt the Code Ocean platform since computational reproducibility is critical to show how you created and executed the results. Code Ocean enables orchestration of existing building blocks, such as toolsets, to get to that level of reproducibility. Code Ocean's well defined Compute Capsule structure ensures standardization for computational research best practices, to easily scale projects across researchers.

With Code Ocean, CytoReason researchers use an intuitive computational workbench that simultaneously supports RStudio, Jupyter, and other popular tools, without the need to setup or install anything locally. They create projects that are organized, accessible, and interoperable, making it easy to reuse and build upon previous work. Internal publishing, sharing, and collaboration are as easy as sending a link and granting permission.

The Code Ocean platform was deployed using CytoReason's Amazon Web Services virtual private cloud to keep projects secure and ensure complete control over data governance. To optimize compute resources and costs, CytoReason now has multiple cloud machines allocated to projects on an on-demand basis, and scaled in by the platform when not in use. Projects start faster and research teams access exactly the needed computing resources, at optimized cost.

Results

Code Ocean enabled fast setup and onboarding of researchers for CytoReason to shorten the learning curve and speed up new project implementation. Over 30 CytoReason researchers transitioned into using the Code Ocean platform in their day-to-day work within a few weeks, creating dozens of projects, and running project computations thousands of times. CytoReason has already seen researcher productivity gains of 15–30%. By automating labor-intensive operational tasks with the Code Ocean platform, CytoReason now has more resources to dedicate to data engineering. They have realized annual savings of more than $100,000 for DevOps work. Efficiencies in compute savings keep paying back over time.

> We have a commitment to quality for our customers, which makes computational reproducibility essential. Code Ocean makes it possible to reproduce projects so we can quickly identify and address quality issues.

Yuval Kalugny, CytoReason VP Engineering

Why Code Ocean

- One integrative computational research experience, emphasizing researchers' productivity and reproducibility
- Highly flexible, configurable, and secure capabilities
- Customer can focus on the discovery logic rather than on tool management
- Best practices with ability to forecast new needs, and deep understanding of market trends
- Superb service, exceeding expectations of operational support, with transparent upgrade processes