

# The Allen Institute implements Code Ocean to streamline workflows

Improves research reproducibility, collaboration, and efficiency

The Allen Institute is a Seattle-based independent nonprofit bioscience and medical research institute founded by Microsoft co-founder and philanthropist Paul G. Allen. It conducts large-scale research through foundational science to fuel the discovery and acceleration of new treatments and cures for diseases such as Alzheimer’s disease, heart disease, cancer, addiction, and more. In keeping with its commitment to open science, one of the Allen Institute’s unique core values is to make all data and resources publicly available for external researchers and institutions to access and use.

## The Background

The Allen Institute is composed of five scientific units and more than 800 employees. Since implementing Code Ocean, the Allen Institute has achieved dramatic improvements in reproducibility, interoperability, and productivity. The organization has been able to complete more at a quicker pace, enabling more than 100 researchers to actively use the platform to make discoveries faster and more efficiently.

Code Ocean’s “no lock-in” platform has also helped the Allen Institute stay true to its mission of advancing open science by reaching and collaborating with more than 200 external users.

To understand the impact of Code Ocean on the organization as a whole, we spoke with Dr. David Feng, Director of Scientific Computing, and Dr. Jerome Lecoq, an Associate Investigator for the Allen Institute for Neural Dynamics (AIND). It’s important to note that Dr. Lecoq also serves as a principal contributor to **OpenScope**, a project that aims to share complex neuronal recording pipelines with the entire neuroscience community. His team’s neuroscience

### Goal

To streamline workflows and to share end-to-end processes with external scientists

### Key Benefits



Increase pipeline building efficiency by 300%



Reallocated 10 team members to higher impact tasks



Improved pipeline interoperability from 25% to 100%

research applies large-scale deep learning, dealing with millions of data points that require extensive processing pipelines to handle and run analyses on complex and rich neuronal datasets. Since the OpenScope project is a highly collaborative and distributed effort, it is of paramount importance that Dr. Lecoq’s team share as much as possible about its data and processes with the broader neuroscience community.

## The Challenge

Many ongoing projects at the Allen Institute, including OpenScope, are using custom pipeline management code running on the on-premise HPC cluster. However, this workflow system was not designed to be exported and shared with outside parties, which is a critical need of the OpenScope project. As a result, when it came to sharing

Dr. Lecoq’s computational process with the broader neuroscience community — a key value of the Allen Institute — he was previously able to share only a small portion of his process, as opposed to the whole end-to-end workflow, because his pipelines were designed to run on the Allen Institute’s specific compute infrastructure.

More generally, Dr. Lecoq found that many laboratories would be eager to adopt processing pipelines validated in dedicated institutions like the Allen Institute, rather than replicate all processes from scratch in their laboratories.

“The neuroscience community is building over and over the same pipelines across hundreds of laboratories,” explained Dr. Lecoq. “So far, we have mostly not been able to leverage the opportunities presented by the Cloud.”

Each one of these pipelines requires dedicated engineers, or more often students that have to learn the entire code stack. It very likely slows our scientific progress.”

Dr. Lecoq, Associate Investigator  
Allen Institute for Neural Dynamics

See how Code Ocean can help you reach your business goals

REQUEST A DEMO



codeocean.com/demo

## The Benefits

To address these challenges, Dr. Lecoq has leveraged Code Ocean’s computational research platform. Once onboarded, Dr. Lecoq began to derive value from the platform within a matter of days. He reports two major benefits to using Code Ocean so far:

### 1 Code Ocean’s “self-service” capabilities.

Code Ocean has enabled Dr. Lecoq and other scientists to independently design and build pipelines, allowing the AIND engineers to shift their focus from routine support tasks to more meaningful projects.

“Code Ocean makes it easy for our scientists to do their work reproducibly,” said Dr. Feng, Director of Scientific Computing, the Allen Institute for Neural Dynamics. “New users to the platform can get far with just a little support... this gives our engineers time to focus on domain-specific challenges.”

2 Increased interoperability within the broader neuroscience community. Using Code Ocean, Dr. Lecoq’s team can share the full pipeline, as opposed to sharing just a portion.

Dr. Lecoq is not the only scientist at the Allen Institute who has benefited from Code Ocean—a few hundred scientists, internally and externally, are also deriving similar benefits from the platform.

Using Code Ocean has resulted in efficiency gains as well as a reallocation of time and resources to higher-leverage activities:

- The speed at which a pipeline can be built has decreased from 12 to 3 weeks — a 4x increase in workflow efficiency.
- The Allen Institute has been able to reallocate the time of 10 engineers to more impactful projects.
- With fully interoperable Code Ocean Pipelines, the time and effort required to share a pipeline has been reduced to just one click.