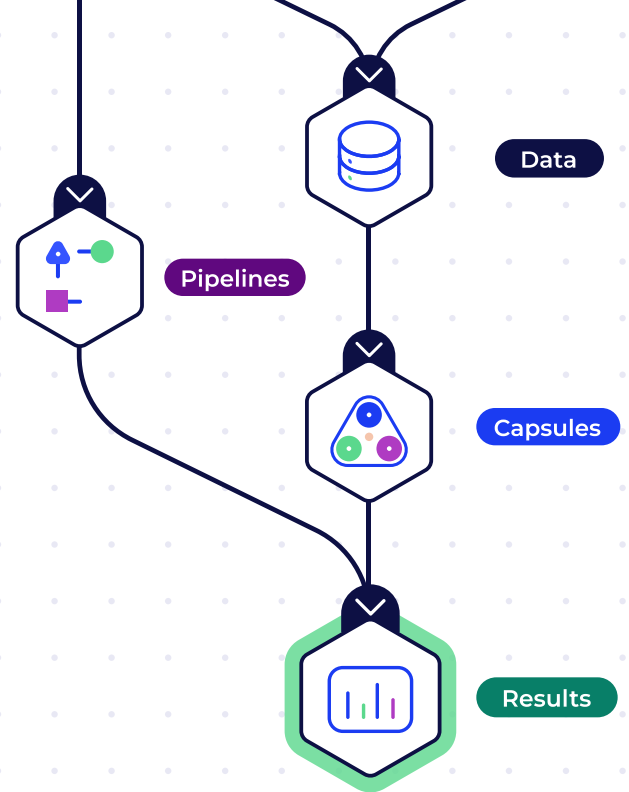# code ocean

Computational Science platform

The best way to run scalable, reproducible computational research

Provision compute, set up environments, develop analysis, scale powerful pipelines, and automatically create a reproducible, traceable record of your work.

Data

Pipelines

Capsules

Results

{X}; Typical use cases

- Multiomics
- Image processing
- Cloud management
- Result provenance
- Data analysis
- Data management
- Bioinformatic pipelines
- ML model training

{◇}; How Code Ocean helps

- Easy to provision, scale, and manage compute resources
- A secure and reliable way for Computational Scientists to work together
- Guarantees complete scientific reproducibility and traceability
- Supports compliance and quality assurance requirements
- Avoids tech debt, and has no lock-in — export anything, at any time

" Code Ocean makes it easy for our scientists to do their work reproducibly ... New users to the platform can get far with just a little support; this gives our engineers time to focus on domain-specific challenges."

Dr. David Feng
Director of Scientific Computing, the Allen Institute for Neural Dynamics

Trusted by Computational Scientists in life science R&D

nature ALLEN INSTITUTE Cajal Neuroscience CytoReason ochrebio Lantern Pharma

Learn more at codeocean.com

code ocean

Data Sources

Data warehouses
Local Files, Databricks, AWS S3 Buckets, GCP Batch, AWS Redshift, snowflake, Collibra

LIMs / ELNs
thermo scientific, Benchling, Dotmatics

Instruments
illumina, thermo scientific, PerkinElmer, 10x Genomics

Apps Library
Code Ocean apps library

HIPAA COMPLIANT
ISO 27001
GDPR
FAIR

No Lock-In Philosophy

Research

Data
- Upload
- Search
- Metadata
- AWS S3
- AWS Athena

Compute Capsules

Compute Environment
- Docker
- General Compute

Code
R, Python, Matlab, C++

Git services
Gitlab, Github, Bitbucket, Azure DevOps

Attach Data
- Results
- External
- Internal
- Data
- Combined Data

Cloud workstations
Jupyter, VS Code, Ubuntu, R Studio, Windows, MATLAB

No code apps
Streamlit, Shiny, IGV

ML Tools
mlflow

Pipelines
Results Bucket, Nextflow, nf-core

Provision

Virtual Private Cloud
- Amazon Web Services
- Amazon Batch
- Amazon EC2
- Amazon S3 Buckets
- Spot Instances

Manage

Admin Panel
- **Permissions & access control** — Azure Active Directory, Okta
- **Analytics and controls**
- **Collaboration**

Data provenance

Lineage Graph
Data, Pipelines, Capsules, Results

Guaranteed reproducibility

Organize and reuse

Collections — computational science library
Capsules, Environments, Results, Data, Pipelines, Apps

Small Molecule, Microbiome, Protein Folding, Proteomics, Pharmaco-genomics, Genomics, Vaccine Efficacy, Biomarker Identification

---

Data Targets

Data warehouses
Local Files, Databricks, AWS S3 Buckets, GCP Batch, AWS Redshift, snowflake, Collibra

LIMs / ELNs
thermo scientific, Benchling, Dotmatics

External workbenches
tableau, Prism, amazon QuickSight, Spotfire

External ML Tools

External Compute
Google Cloud Platform, Microsoft Azure, On Prem

---

codeocean.com