

# NIH STRIDES AND AI-ENABLING NIH RESOURCES

---

**Developed for AIM AHEAD**

NIH STRIDES Initiative  
Center for Information Technology, National Institutes of Health

May 2024

# Opportunities & Challenges of the Cloud

Cloud is the new frontier of biomedical research with ample benefits and some barriers.

## THE OPPORTUNITIES



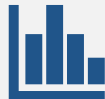
**Simplifies testing** and iteration



**Democratizes** access to scientific data



**Always available** on demand



**Easily scales** up and down



**Pay as you go** for only what you use



**Provides a rich set** of tools & services

## THE CHALLENGES



**Acquisition** is complex and time-consuming



**Security protections** are unclear



**Costs** are unpredictable



**Training deficit** is substantial

NIH offers two programs that help funding recipients **address those challenges.**

## NIH STRIDES Initiative

- Purchase long-term commercial cloud services
- Purchase training from STRIDES Cloud Service Provider partners
- Purchase Professional Services support from STRIDES Cloud Service Provider Partners

## NIH Cloud Lab

- Temporarily try the cloud at no cost to test your use case

# NIH STRIDES Initiative

Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability

## Overview

The STRIDES Initiative serves **both the NIH Intramural and Extramural Research Communities** and accelerates biomedical research in the cloud by:

- Simplifying access
- Reducing costs
- Lowering technological barriers
- Improving processes

## Core Motivations

- Democratization of computational research and data science
- Cost savings and efficiencies for the research community
- Strong partnerships with cloud providers



Partnerships with



Google Cloud



Microsoft Azure

NIH STRIDES OVERVIEW

# Cloud Training at NIH

---

# Cloud Training Courses

The NIH STRIDES Initiative enables NIH staff and NIH-funded researchers to access group training from Amazon Web Services, Google Cloud, and Microsoft Azure **at a lower rate.**

## Training Formats



### Instructor-Led

Live, scheduled **Virtual** or **In-Person** courses led by instructors



### Self-Paced

Self guided e-learning that staff can access around their work schedule

## How To Access Training

- Browse the STRIDES Training homepage to explore all available courses
- Browse through the AWS, Microsoft Azure, and Google Cloud course catalogues and consult your preferred training provider to purchase at NIH lower rate:

[AWS Course Catalogue](#)



[Azure Course Catalogue](#)



[Google Cloud Course Catalogue](#)



# Custom Course Series

Medical Imaging on Google Cloud Course Description



STRIDES CUSTOM COURSE SERIES



## Medical Imaging Suite and Image Machine Learning on Google Cloud

**1 Day, Instructor-Led Virtual**

In this one-day, virtual course, you'll learn how to use [Google Cloud Medical Imaging Suite](#), AI-powered tools, and various cloud-based applications to process images and perform advanced image analysis. This course includes guided, hands-on labs where you'll learn directly in a cloud environment and use Imaging Data Commons (IDC) data.

This course is part of a series of custom courses that [STRIDES](#), NIH's cloud adoption initiative, developed in partnership with commercial cloud service providers to meet the unique training needs of NIH's scientific community. To learn more about other custom trainings, [visit the STRIDES Custom Course Series SharePoint folder](#).

### Who Should Take This Course

This training is intended for biomedical computational researchers and research IT staff who work with imaging systems such as cameras, microscopes, MRI, CT, and ultrasound.

### What You'll Learn

#### Getting Started with Google Cloud

This topic will introduce you to the Google Cloud environment, including the console, [Cloud Shell](#), and Cloud Shell Editor. You'll learn about the [Cloud SDK](#), a set of tools for interacting with Google Cloud products and services, and [Google Cloud Skills Boost \(Owikilabs\)](#), a

STRIDES developed custom course series in partnership with Google Cloud and Amazon Web Services to meet the unique training needs of NIH's scientific community.

# NIH Cloud Lab

---

# NIH Cloud Lab | Experiment in the Cloud

NIH Cloud Lab is a no-cost, 90-day program for researchers to **try Amazon Web Services, Google Cloud, or Microsoft Azure** in an NIH-approved environment.

The team launched Cloud Lab to NIH staff in June 2022 and a small group of awardees in June 2023. We're now expanding the program to current and prospective recipients of NIH funding.



**Request an Account or Learn More!**

Scan the QR code or visit  
[cloud.nih.gov/resources/cloudlab](https://cloud.nih.gov/resources/cloudlab)

## NIH Use Cases

---

### Evaluate Utility & Cost

Provides an easy route to evaluate the cloud's utility/cost for a project without major time or financial commitments

### Develop New Tools

Allows experienced teams to prototype new architectures and evaluate software and hardware combinations

### Share Ideas

Connects NIH'ers from across ICs to share ideas on how to conduct biomedical research in the cloud

### Learn New Skills

Simplifies access to tools and cloud environments that participants can use for training purposes



# How It Works

Cloud Lab is designed to be a widely accessible, fully customizable, and NIH-managed training tool, **not a production research environment or a place to work with sensitive data.**

## Services help trainees grow their cloud skills



### Account and Credits

90 days of access plus \$500 of cloud credits that participants can use to deploy a full range of resources.



### Tutorials & Data Sets

Access to tutorials that demonstrate how to run real AI, bioinformatics, and data science workflows.



### Support

A dedicated mailbox and one-on-one support from a team of technical and bioinformatics NIH experts.



### Weekly Tips Emails

Emails with information on how to get started, helpful resources, and weekly tips.

## Seamless Cloud Access

### 1. REQUEST AN ACCOUNT

...by completing a form on [Cloud Lab Page](#)

### 2. GET A CLOUD ACCOUNT

...in as little as few business days

### 3. LOG IN

...to access \$500 of credits for 90 days

### 4. LEARN & EXPLORE

...using resources provided by Cloud Lab

# Interactive Tutorials

Cloud Lab participants can use their accounts to run interactive, cloud-based tutorials in several GitHub repositories.



## NIH Cloud Lab for Amazon Web Services

This repository teaches Cloud Lab participants and others how to use Amazon Web Services for research. It includes basic tutorials and links to helpful resources.



## NIH Cloud Lab for Google Cloud

This repository teaches Cloud Lab participants and others how to use Google Cloud for research. It includes basic tutorials and links to helpful resources.

## NIGMS Sandbox

This repository teaches students, researchers, and clinicians how to utilize the cloud for research. It includes 12 interactive modules and a set of [complementary videos](#).



## NIH Cloud Lab for Azure

This repository teaches Cloud Lab participants and others how to use Azure for research. It includes basic tutorials and links to helpful resources.

**Note:** These tutorials are fully public and available for anyone to use with a self-funded cloud account.

# Rich Tutorial Content

Cloud Lab offers a **suite of interactive tutorials** designed to help participants perform viable research in the cloud.

## Scientific

- Accessing Open Datasets
- AI/ML
- ATACseq (NIGMS)
- Biomarker Discovery (NIGMS)
- BLAST
- FHIR
- Genomic Pathway Analysis (NIGMS)
- Genomic Variant Calling
- GWAS
- Long Read Genomics
- Medical Image Analysis (NIGMS)
- Medical Image Segmentation
- Metagenomics (NIGMS)
- Proteomics
- RNAseq
- scRNAseq
- Radiogenomics
- SRA Data Interaction
- VCF Query with SQL
- ...and more

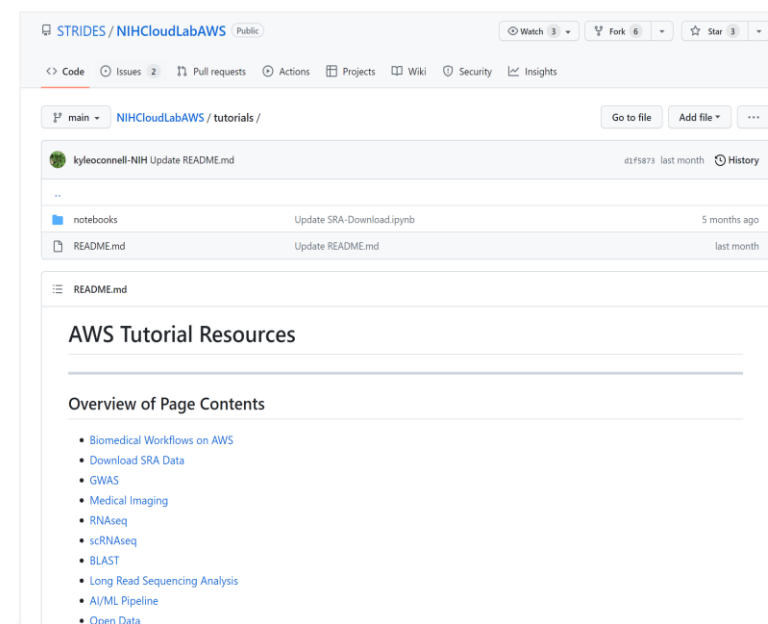
## General Cloud Ops

- Access Marketplace Offerings
- Access Public Datasets
- Command Line Tools / SDKs
- Disk Images
- Ingest and Store Data
- Introduction to Cloud
- Jupyter Notebooks
- Serverless Functionality
- Virtual Machines

## Computing & Code

- Conda Environments
- Container Registries
- Git Repos
- HPC Clusters
- Kubeflow
- Serverless Functionality

### NIH Cloud Lab AWS Tutorial Repository



# Generative AI at NIH

---

# NIH Generative AI Community of Practice (GenAI CoP)



## Vision

- Harness the potential of GenAI across NIH to enable biomedical research discoveries and accelerate NIH business operations



## Goals

- Serve as the nexus for GenAI coordination, innovation, and deployment
- Encourage, test, support, and scale GenAI adoption at NIH



## Objectives

- Foster collaboration and encourage cross-disciplinary interactions
- Share knowledge and ideas via seminars, workshops, forums, etc.
- Upskill & develop staff
- Promote responsible and secure use of AI
- Adapt to the community's needs and establish channels for cross-IC support

# Current State of GENAI at NIH

## KNOWN NIH USE CASES



30+

## TOTAL CoP MEMBERS



600+

### NIH GenAI CoP is quickly growing and maturing at NIH

- Since January 2024, we held 5 CoP meetings, averaging 200 attendees per event

### GenAI Services being delivered rapidly to support the NIH community via GenAI Process and Solutions Portfolio

- Guide use case development, host technical talks, and share CSP capabilities
- Use cases include Customer Response, Metadata Transformation, Classification/Accessibility, Data visualization, IT Programming Tasks, Research & Development

### Accounts and Tech Support

- Create STRIDES accounts or subscriptions, ensure access to data sources, create OpenAI instances, and manage underlying policy and security guardrails

### GenAI Tutorials

- Developed tutorials ranging from beginner to advance levels across CSPs for the STRIDES GitHub

# Supporting GENAI Users

## STRATEGY & ENGAGEMENT

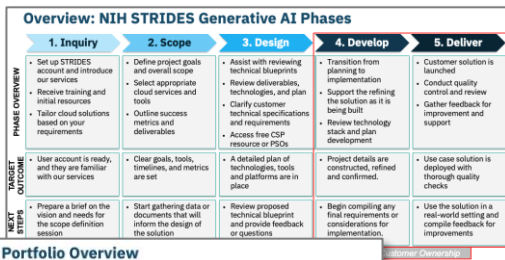
- Launched **GenAI Process** for customer transparency
- **GenAI Solutions Portfolio** to highlight CSP capabilities

## TECH & OPERATIONS

- Provide GenAI support across 3 CSP Partners
- Use Service Now to track GenAI consult and support requests

## CONSULTS & TUTORIALS

- Developed **25+ GenAI tutorials** and How-To docs
- Provide user tutorial support via **1-1 sessions**



The Requester Details for NIH users come from the NIH Employee Directory (NED). You can go to NED to update your record if it is inaccurate. For HRSA users, go to ADM (Active Directory Management). If you are submitting this request for someone else, please change the Requested for field below. To provide alternate contact information only for this ticket, such as a different phone number if you are on travel or teleworking, please update the Phone Number field below:

\* Requested for:

Long Org/Division:

Email:

Building/Location:

Phone:

**Cloud Services**

- ✓ Cloud Software
  - NIH GitHub Request
- ✓ Enterprise Cloud Platforms
  - Cloud Consultation Requests
  - Cloud Operations Support Request
  - New Cloud Account Requests

Azure AI services | Azure OpenAI

Search:

+ Create Manage deleted resources Manage view Refresh Ex

Filter for any field... Subscription equals

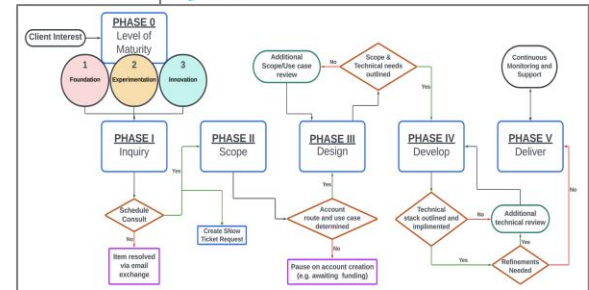
All Azure AI services Showing 1 to 1 of 1 records.

Azure AI services	Name	Kind	Location
Azure OpenAI	nihs-openai-poc	OpenAI	East US

**GenAI Solutions Portfolio Overview**

The [NIH STRIDES GenAI Solutions Portfolio](#) provides an introduction to the GenAI capabilities within the NIH STRIDES program which empowers researchers with direct access to the latest models and custom use cases via our industry-leading cloud platform partners: Amazon Web Services (AWS), Microsoft Azure with OpenAI (Azure), and Google Cloud Platform (GCP).

The program collaborates closely with users to ensure seamless integration and tailored experiences, enabling them to leverage the full potential of Generative AI while offering secure, NIH-approved environments to explore and learn about cloud computing & AI.



# GitHub Copilot

Increase developer productivity with GitHub Copilot.



## GitHub Copilot

### About GitHub Copilot

GitHub Copilot is an AI-powered coding assistant that leverages GenAI models to provide intelligent code suggestions and automate repetitive coding tasks. It is trained to use code from public repositories, allowing it to understand context and generate relevant code snippets on the fly.

Some research points to productivity gains of up to 40% when using GitHub Copilot across the entire development lifecycle. At NIH, Enterprise GitHub, Copilot is available for all the Organizations. For access and details, contact [github@nih.gov](mailto:github@nih.gov).



# Demo

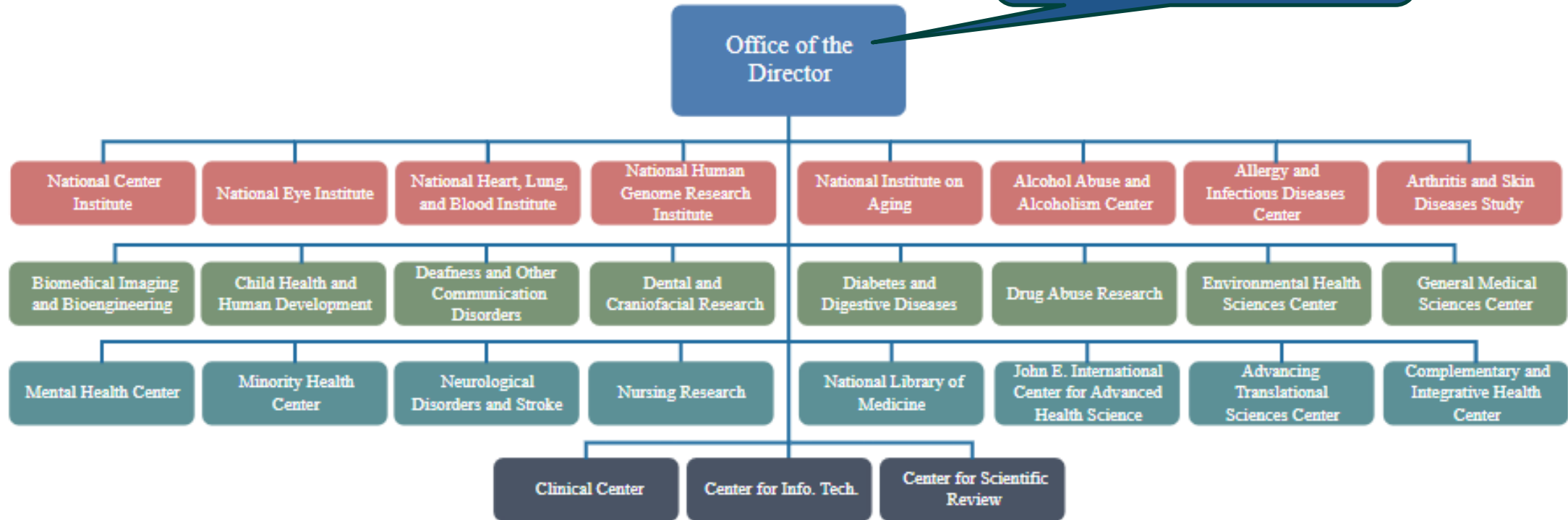
NIH Cloud Lab & GenAI

# Additional NIH Resources

---

# NIH and the Office of Data Science Strategy (ODSS)

Office of Data Science Strategy (ODSS)



# NOSI: Promoting Data Reuse for Health Research

- The NIH announces a new NOSI for competitive revision applications focused on data reuse and secondary data analysis in NIH-funded data repositories to address health research questions.
- Aims to benefit early-career researchers and those at resource-limited institutions through improved data availability and documentation.
- Strengthens NIH's data science infrastructure by fostering data sharing and adhering to FAIR principles (Findable, Accessible, Interoperable, Reusable).

**First Available Due Date:** July 03, 2024



**Learn more here:**  
<https://bit.ly/3V7rd6D>

# NOSI: Advance Data Science Approaches Through Secondary Data Analysis to Reveal Scientific Insights of COVID-19 Testing Technologies

- Stimulating data science approaches related to COVID-19 testing technologies.
- Advancing scientific inquiry through secondary analysis of existing data in the RADx DataHub.
- Addressing questions and revealing insights through secondary analysis.
- Focusing on utilizing existing data resources in the RADx DataHub.

**Expiration Date:** July 16, 2024



**Read the NOSI:** <https://bit.ly/4a4oFf3>

# RECOVER (Long COVID) Data Available in BDC

- RECOVER researchers collect data from numerous study participants, making it available to the public and scientific community for improved collaboration.
- Those interested can access aggregated study data without approval, while de-identified individual data requires authorization.
- Authorized researchers access RECOVER data through BioData Catalyst® (BDC) and publication-specific portals, as listed on the RECOVER publications page.



Learn more here:  
<https://bit.ly/3wc9OAK>



# Medical Image De-Identification Benchmark (MIDI-B) Challenge

- The MIDI-B Challenge invites developers to assess their DICOM image de-identification tools, focusing on automated methods that preserve research utility.
- Participants will use diverse medical imaging data from NCI's Cancer Imaging Archive and can co-author the MIDI-B Challenge report.
- Top teams will present their algorithms and results at the MICCAI 2024 virtual session, promoting advancements in medical imaging and AI.

**Registration Ends:** August 15, 2024



**Learn more here:**  
<https://bit.ly/3UMhGRi>

# New Funding Opportunity: Collaborative Research Network

- Join ODSS in a collaborative research network to test preventive interventions for health disparities.
- Funding Opportunity: PAR-24-053 encourages multi-sectoral projects beyond individual health.
- Funding Opportunity: RFA-OD-24-006 seeks collaboration for technical assistance in community-focused projects.

**Application deadline:** Aug. 5, 2024



**Apply for PAR-24-053 here:**  
<https://bit.ly/3ujn5pJ>



**Apply for RFA-OD-24-006 here:**  
<https://bit.ly/3HMNz5X>



# Basic Data Science Training Resources Available From NCI

- New cancer data science training resources provide resources that cater to both newcomers and seasoned researchers in the field.
- Encourages individuals to explore basic resources, free courses, tips, etc.
- Help leverage the power of data science for your cancer research.



**NATIONAL CANCER INSTITUTE**  
Center for Biomedical Informatics  
& Information Technology



**Check out the new resources here:**

<https://bit.ly/46Jv3FJ>

# Introducing the NCI Computational Resources for Cancer Research Portal

- Researchers can access cancer-related software, datasets, and AI models through the Computational Resources for Cancer Research portal, including those from the NCI-DOE Collaboration.
- The portal features use cases, educational materials, emerging areas updates, and event notices aimed at fostering cross-disciplinary collaboration among researchers.
- By joining the community, researchers contribute to the portal's growth and help impact cancer research.



Learn more here:

<https://bit.ly/4b71EYU>

# ODSS Funding Announcements

## ODSS Priority Areas\*:

- Integrated Infrastructure and Emerging Technology (IIET)
- FAIR Data & Resources
- Clinical Informatics
- Training, Workforce Initiatives and Community (TWICE)
- COVID & Policy
- AIM-AHEAD - Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity (AIM-AHEAD)



Harnessing Artificial Intelligence and Polypharmacology to Discover Pharmacotherapeutics for Substance Use Disorders (R41/R42 Clinical Trials Not Allowed) | NIDA

- Link: [RFA-DA-25-053](#)
- Submission date July 25, 2024

Four NCI NOFOs on Informatics Technologies for Cancer Research and Management | NCI

- Links: [RFA-CA-24-016](#), [RFA-CA-24-017](#), [RFA-CA-24-018](#), [RFA-CA-24-019](#)
- Submission date November 15, 2024

Smart Health and Biomedical Research in the Era of Artificial Intelligence and Advanced Data Science (SCH) | NIDA

- Links: [NSF 23-614](#) [NOT-OD-23-165](#)
- Submission dates October 03, 2024; October 3, 2025

Notice of Special Interest: Leveraging Existing and Accessible Datasets for Implementation Research Strategies and Testing - LEAD FIRST | NHLBI

- Link: [NOT-HL-23-120](#)
- Standard submission dates Feb. 5, Jun. 5, Oct. 5 until Feb. 5, 2025

Personal Health Informatics for Delivering Actionable Insights to Individuals (R01 Clinical Trial Optional) | NLM/NIMH

- Link: [PAR-23-245](#)
- Standard submission dates Feb. 5, Jun. 5, Oct. 5 until Oct. 5, 2026

\*Request for Information: Inviting comments on the NIH Strategic Plan for Data Science 2023-2028 (closed 03/05/24)

# Thank You!

---