



# **AIM-AHEAD *All of Us* Training Program**

## **Cohort 3**

**Informational Webinar**

**September 11, 2025, 4:00–5:00 PM ET**

# Before we get started...



- This session is being recorded. The recording and presentation slides will be shared after the webinar.
- In an effort to reduce accidental unmuting and background noise, we have disabled attendee microphones until the Q&A portion.
- We will hold a Q&A session at the end. You may drop your questions in the chat or electronically “raise your hand” at that time.

*Thank you for joining!*

# About AIM-AHEAD Consortium



*AIM-AHEAD is a nationwide network of institutions and organizations designed to build AI talent among researchers and clinicians, support multidisciplinary research projects that harness AI/ML to improve the health of Americans and enhance the AI capabilities and infrastructure of communities or hospitals that otherwise would not have had the resources or the capacity to benefit from the advance of AI/ML.*

# The AIM-AHEAD Coordinating Center



The A-CC consists of four cores, focused on various initiatives to achieve AIM-AHEAD's mission.

## Leadership Core

Lead, recruit, and coordinate the AIM-AHEAD Consortium

## Data Science Training Core

Assess, develop, and implement data science training curriculum

## Data and Research Core

Address research priorities and needs to form a comprehensive basis for AI/ML

## Infrastructure Core

Assess data, computing, and software infrastructure to facilitate AI/ML and health research





# NIH Leadership Team



**Samson Gebreab, Ph.D. MSc.**

*Program Lead, AIM-AHEAD  
Office of Data Science Strategy,  
NIH*



**Eva Lancaster, Ph.D.**

*Program Director, AIM-AHEAD  
Office of Data Science Strategy,  
NIH*



**Emir Khatipov, Ph.D.**

*Program Director, AIM-AHEAD  
Office of Data Science Strategy,  
NIH*



**Sydney J. McMaster, MPH**

*Research Engagement and  
Outreach Team Lead, Division of  
Engagement and Outreach, All of Us  
Research Program, NIH*

# Meet the AIM-AHEAD Team

**All of Us**  
RESEARCH PROGRAM

**IRT**  
INTERNATION



**Toufeeq A. Syed, Ph.D.**  
UTHealth Houston  
*All of Us* Co-Director and  
AIM-AHEAD MPI



**Legand L. Burge, III, Ph.D.**  
Howard University  
*All of Us* Co-Director and  
AIM-AHEAD MPI



**Robert T. Mallet, Ph.D.**  
UNT Health Fort Worth  
*All of Us* Co-Director



**Peter McGarvey, Ph.D.**  
Georgetown University  
AIM-AHEAD MPI, DSTC Lead -  
Foundations of Data Science  
for Health care Applications



**Lesley Grandstaff, M.Ed.**  
*All of Us* Program Manager,  
Communications Hub



**Alyssa Parham**  
*All of Us* Program Manager,  
Data Science Training Core



**Sarah Popal**  
*All of Us* Program Coordinator,  
Communications Hub



# Meet the RTI Team

**All of Us**  
RESEARCH PROGRAM

**RT**  
INTERNATION



**Megan Lewis, Ph.D.**  
Principal Investigator



**Javan Carter, Ph.D.**  
Lead Genomics  
Coach and Instructor



**Edward Preble, Ph.D.**  
Lead ML Coach and  
Instructor



**Jonathan Holt, Ph.D.**  
ML Coach and  
Instructor



**Sheri Cates**  
Program Manager

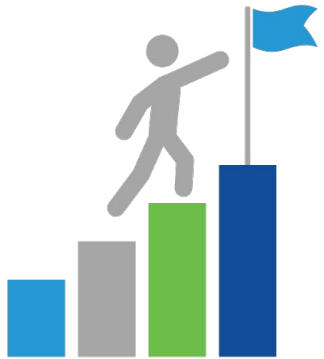


**Leslie Zapata Leiva**  
Program Coordinator



**Ife Adewumi**  
Program Coordinator

# Program Purpose



The AIM-AHEAD *All of Us* training program is intended to expand research in AI/ML by leveraging the **All of Us data and infrastructure (Researcher Workbench)**. To achieve this, the program trains individuals from various multidisciplinary backgrounds to gain proficiency in AI/ML data analysis and apply their expertise to accelerate health research and improve the health of all communities.



# About *All of Us* Training Program



- To accomplish this goal, selected trainees will complete an **8-month training program in advanced data analysis** developed by RTI International (RTI) and utilizing the resources of the *All of Us* database and AIM-AHEAD's Data Science Training Core. Trainees will be equipped to **conduct in-depth analysis of large datasets** essential for cutting-edge biomedical and AI research.
- Trainees will
  - (1) **apply R or Python** to analyze *All of Us* data
  - (2) **formulate hypotheses** testable by applying AI/ML and advanced data analysis
  - (3) **present their work** at the 2026 AIM-AHEAD annual meeting.

# All of Us Data Set

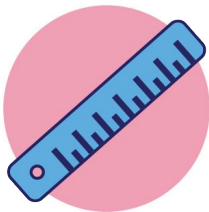


## The *All of Us* data set securely captures a range of different data types

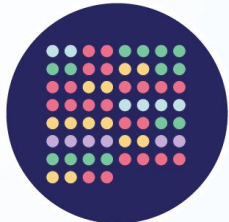
The *All of Us* Research Program's Data and Research Center curates a range of different data types as part of the data collection process. The numbers below reflect the number of participants with each data type available.



**>633,000**  
with survey  
responses



**>509,000**  
with physical  
measurements



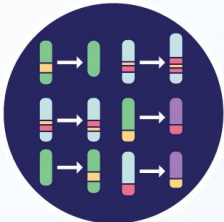
**>447,000**  
with genotyping  
arrays



**>414,000**  
with short-read whole  
genome sequences  
(WGS)



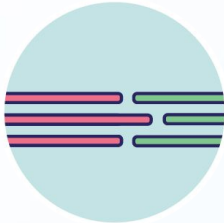
**>393,000**  
with electronic  
health record data



**>97,000**  
with structural  
variant data



**>59,000**  
with Fitbit  
records



**>2,700**  
with long-read  
WGS

Data as of February 2025 Curated Data Repository (CDR) v8 data release

The *All of Us* data set is accessible to researchers across stages and settings

**All of Us**  
RESEARCH PROGRAM

U.S. RT  
INTERNATION



## Our Researchers



**>18,300**

Registered  
Researchers



**>1,100**

Academic  
Institutions



**>170**

Not-for-Profit  
Organizations



**>130**

Health Care  
Organizations

## Research Currently Underway



**>19,500**

Active projects



**>700**

Publications in  
peer-reviewed  
journals



Top conditions being studied in the  
Researcher Workbench include:

- Cancer
- Cardiovascular disease
- Diabetes
- Mental health
- Hypertension

Figures accurate as of August 2025



# Program Expectations for Trainees



Onboarding	Complete the onboarding process including attending all orientation sessions
Time Commitment	Work on the program an average of at least 8 hours per week
Courses	Attend and participate in all the required courses and data use case training for the trainees' tier (Fundamental, Advanced, Genomics Track)
Networking	Engage in AIM-AHEAD learning communities and the peer networking sessions
Mentorship	Engage with an AIM-AHEAD career/professional development mentor (to be assigned through the program)
Assignments	Complete all the assigned goals and tasks in AIM-AHEAD Connect
Research Project	Conduct a research project using <i>All of Us</i> data and present a work-in-progress research poster at the AIM-AHEAD Annual Meeting (July 2026). The project submitted—including the title—must comply with the <a href="#">US White House Executive Order</a> .
Program Evaluation	Participate in program evaluation activities
Engagement	Take an active role in the AIM-AHEAD network by engaging in upcoming events, webinars, and office hours; seeking additional mentorship; and applying to current and future programs and Calls for Proposals

# NEW Genomics Track!



<b>Description</b>	This specialized track is designed to support trainees interested in conducting genomic research using complex ML tools. The track dedicates more time and resources to developing foundational and applied skills in genomic data science.
<b>Selection</b>	Five trainees will be selected. Applicants not selected will still be considered for the General Track.
<b>Additional Criteria</b>	<ul style="list-style-type: none"><li>• Demonstrated proficiency in R or Python</li><li>• General understanding of genetics and genomics related to the applicants' field of study (i.e., undergraduate course work or practical experience)</li><li>• Prior experience in genomics computational applications related to the applicants' field of study</li></ul>
<b>Additional Expectations</b>	<ul style="list-style-type: none"><li>• Participate in the three genomics courses and associated activities (up to 9 additional hours).</li><li>• Meet regularly with the trainee's assigned genomics coach during biweekly group office hours.</li></ul>

# Timeline for Training and Research Project

**All of Us**  
RESEARCH PROGRAM

**IRT**  
INTERNATION





# Overview of Training Curriculum



	Fundamental Tier	Advanced Tier	Genomics Track
Researcher Workbench	How to use the Workbench (2 courses)		
AI Ethics	All of Us and Responsible AI		
Hands-on Case Studies	Data use case training for <b>fundamental</b> tier	Data use case training for <b>advanced</b> tier	Data use case training for <b>advanced</b> tier
Required Technical Coursework	Asynchronous training on either <b>R</b> or <b>Python</b> (trainee selects preferred language)	<b>No required technical coursework for this track</b>	<ul style="list-style-type: none"> <li>Genomics 101</li> <li>Genomics Data Types and Utilities</li> <li>All of Us ML Genomics Analyses</li> </ul>
Optional Technical Coursework	<ul style="list-style-type: none"> <li>AIM-AHEAD Python Bootcamp or R Bootcamp (if needed)</li> <li>Statistical Techniques and Machine Learning for Research Analysis</li> <li>Advanced tier data use case training</li> </ul>	<ul style="list-style-type: none"> <li>Statistical Techniques and Machine Learning for Research Analysis</li> <li>Fundamental tier data use case training</li> </ul>	<ul style="list-style-type: none"> <li>Statistical Techniques and Machine Learning for Research Analysis</li> <li>Fundamental data use case training</li> </ul>

# Program Benefits



## Stipend

An \$8,000 stipend upon successful completion of trainee milestones

Travel allowance to attend the AIM-AHEAD Annual Meeting in summer 2026



## Support

Support and guidance from an experienced AIM-AHEAD mentor

Support from the AIM-AHEAD Data Science Training Core

Direct 1:1 guidance, virtual office hours, HelpDesk support, and concierge services supporting projects using *All of Us* data



## Training

Training on:

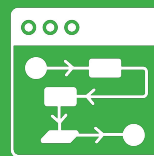
- Using the *All of Us* Researcher Workbench to create a cohort
- Programming in R or Python and Jupyter Notebook
- Merging and validating data across *All of Us* data sources
- Building a supervised model
- Splitting data into subsets for model training and testing, considering anomalies that may be present and undetected or missed by the model, and validating the model
- *All of Us* and Responsible AI
- Genomics track - specialized training on conducting genomics analysis for *All of Us* data

# AIM-AHEAD Mentorship Process



Each trainee will be matched with a mentor who will provide ongoing support throughout the training program. Mentors are matched with mentees using the Connect Platform. Mentorship matches are made using:

**AIM-AHEAD  
CONNECT**



**AI Algorithm**



**Administrative  
Matching**



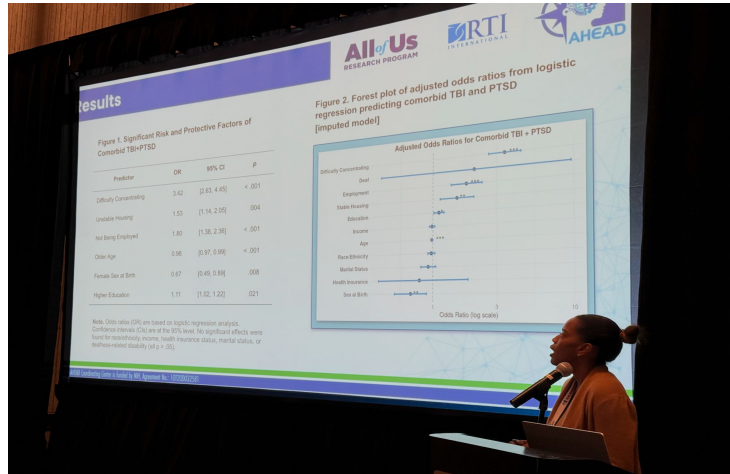
**Mentor Pool  
Search**



# AIM-AHEAD Annual Meeting

**All of Us**  
RESEARCH PROGRAM

**RTI**  
INTERNATION



**Attend engaging panels and presentations**

**Present a work-in-progress poster of research**

**Network with other trainees, awardees, and leadership**

# Applicant Eligibility

## Citizenship and Tax Requirements

- ✓ Must be a U.S. Citizen, Permanent Resident, or Non-Citizen U.S. National
- ✓ Permanent Residency must be established by Sept. 26, 2025
- ✓ Temporary visa holders (F1, J1, H1, etc.) are not eligible
- ✓ Accepted candidates must be able to submit a W-9 form

## Participation Restrictions

- ✓ Current/former AIM-AHEAD program participants (awardees, fellows, trainees, mentors, advisors, coaches) are ineligible
- ✓ Applicants may apply to more than one AIM-AHEAD training program but can only be selected for one

## Special Cases

- ✓ AIM-AHEAD Coordinating Center personnel and Federal employees may participate, but will not receive stipend or travel allowance

*Note: Please refer to the full CFA on [AIM-AHEAD.net](http://AIM-AHEAD.net) for all eligibility requirements.*

# Education & Experience – General Track

## Education Requirements

- ✓ Minimum: Bachelor's degree in physical sciences, life sciences, math, statistics, data science, engineering, health sciences, or public health
- ✓ Eligible applicants include: post-baccalaureate and graduate students, postdocs, medical students/residents, allied health trainees, early-career investigators, and early-career professionals in non-academic institutions

## Required Skills

- ✓ Prior programming experience
- ✓ Basic understanding of statistics
- ✓ Working command of English

## Recommended Skills

- ✓ Coursework in probability and statistics or higher-level math
- ✓ Coding experience in R or Python
- ✓ Experience with data manipulation and management through coursework or research

*Note: Please refer to the full CFA on [AIM-AHEAD.net](http://AIM-AHEAD.net) for all eligibility requirements.*



# Additional Requirements for Genomics Track

## Required Skills

- ✓ Demonstrated proficiency in R or Python
- ✓ General understanding of genetics and genomics as it relates to the applicant's field of study (i.e., undergraduate course work or practical experience)
- ✓ Prior experience in genomics computational applications as it relates to the applicants' field of study

*Note: Please refer to the full CFA on [AIM-AHEAD.net](http://AIM-AHEAD.net) for all eligibility requirements.*

# Eligibility Requirements

## Eligible Organizations

**Higher education institutions**

**Local, state, and tribal governments**

**Nonprofits**

(with or without 501(c)(3) status)

**Other U.S.-based organizations**

(e.g., school districts, housing authorities, faith-based, community-based, and regional organizations)

**For-profit businesses and organizations**

## DURA Requirement

In order to gain access to the Controlled Tier your institution must have a DURA in place

1. [Check here](#) if your institution already has a DURA
2. If not, **begin now** by contacting your institution's office of research/sponsored programs. For more information visit: [How to Obtain a DURA with All of Us – User Support](#)

*Note: Please refer to the full CFA on AIM-AHEAD.net for all eligibility requirements.*

# Application Requirements



**Submission Deadline:** September 26, 2025 by 11:59 PM EST

## Required Application Elements

- **Profile Information** (basic applicant details in InfoReady portal)
- **Letters of Support & Recommendation**
  - Supervisor letter confirming protected time
  - At least one faculty recommendation (additional letters optional)
- **Academic Transcript** (undergraduate and, if applicable, graduate)
- **Biographical Sketch** (NIH biosketch or CV, max 5 pages)
- **Statement of Rationale** ( $\leq 2$  pages) describing goals, need for training, research question, relevant experience, and long-term plans
  - **Genomics Track Statement of Rationale** is separate

## Important Note

- Applicants may apply to more than one program but can only be selected for one
- Applicants will rank program preferences in the application

*\*This is just an overview. Please see the CFA for the full list of application requirements*

# How to Apply

***Applications must be submitted between September 02, 2025 and September 26, 2025 at 11:59 PM EST***

**Note:** Please use Chrome, Firefox, or Edge browsers.

1

**Familiarize  
yourself with the  
program  
requirements  
outlined in the call  
for applications**

2

**Gather all of  
the required  
application  
materials**

3

**Create an account  
on AIM-AHEAD  
Connect and  
register as a  
"mentee/learner"**

4

**Submit  
application for  
review using the  
InfoReady  
platform**



**Up to 25 trainees will be selected**



# Key Program Dates



**CFA Release Date**

September 02, 2025



**Application Deadline**

**September 26, 2025 by 11:59 PM EST**



**Notice of Award**

November 10, 2025



**Program Start Date**

**November 17, 2025**



**Deadline to obtain DURA**

**December 15, 2025**



**AIM-AHEAD Annual Meeting 2025**

July 2026



**Program End Date**

**July 31, 2026**

# Questions?



Please see the FAQ document linked above and in the chat



Scan the QR code above to access the AIM-AHEAD *All of Us* CFA.