



Artificial Intelligence/**M**achine Learning Consortium
to Advance **H**ealth **E**quity **A**nd Researcher **D**iversity

SYLLABUS: PRIME COHORT OF 2022

AIM-AHEAD DATA SCIENCE TRAINING CORE TRAINING PRACTICUM PROGRAM

PROGRAM LEADERSHIP

Director: **Nawar Shara, PhD**, Nawar.Shara@MedStar.net

co-Director: **Alexander Libin, PhD** Alexander.Libin@MedStar.net

Curriculum Development: **Omar Aljawfi, PhD** omar.m.aljawfi@medstar.net

Program Managers:

Sara Stienecker Sara.L.Stienecker@MedStar.net, **Tom Whitfield** Tom.Whitfield@Medstar.net, & **Alyssa Parham** Alyssa.Parham@Howard.edu

The help Desk is open for questions. If you must send a direct email, please direct your curriculum questions to Dr. Aljawfi, programmatic questions to Sara Stienecker, and financial questions to Alyssa Parham.

PRIME GOALS

- AIM-AHEAD's Mission:

The National Institutes of Health's Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity (AIM-AHEAD) program has established mutually beneficial, coordinated, and trusted partnerships to enhance the participation and representation of researchers and communities currently underrepresented in the development of AI/ML models and to improve the capabilities of this emerging technology, beginning with electronic health records (EHR) and extending to other diverse data to address health disparities and inequities.

- Pilot seed awards intended for students to engage in real-life **hands-on training projects** using AI/ML, big data, EHR, and Social & Environmental Determinants of Health (SEDoH) to **address issues in health disparities**
- Learn how to build **equitable and just AI/ML applications** for both professionals and communities with under-represented racial and ethnic identities.

SYNCHRONOUS MEETINGS

There will be a few meetings throughout your engagement with PRIME that require you to attend “live.” These meetings will be recorded and available for those who reach out to Dr. Libin and Sara Stienecker in advance. Your active participation in these meetings is expected as part of participating in PRIME and should also be recorded on your weekly hourly tracking and attestation form.

Date	Meeting	Goals	Outcomes
09/09/2022 12:00 pm – 1:00 pm EST	Kickoff Meeting	<ul style="list-style-type: none"> Meet program leadership Introduction to program setup Q+A 	<ul style="list-style-type: none"> Identify who to contact for your questions Finalize all the paperwork for all the trainees.
09/20/2022 2:00 pm – 3:00 pm EST	Kickoff Meeting 2	<ul style="list-style-type: none"> Trainee introductions Program curriculum overview 	<ul style="list-style-type: none"> Get more details about the trainee’s interest in research Determine elective courses applied into the program
Date TBD – mid-October Time TBD	Mentor/Mentee/Admin meeting	<ul style="list-style-type: none"> Introduction Expectations Roles for home institution mentor & AIM-AHEAD mentors Networking 	<ul style="list-style-type: none"> Match each trainee with an internal mentor through the AIM-AHEAD platform.
Date TBD TIME TBD (1 hr)	AIHES2022 Workshop Series kickoff	<ul style="list-style-type: none"> Big picture introduction to AIM-AHEAD and AI/ML principles in healthcare & health equity 	<ul style="list-style-type: none"> Watch all the workshop series 12 videos.
Date TBD TIME TBD (1 hr)	DataQuest Hands-on Practice Module kickoff	<ul style="list-style-type: none"> Access to the training trainees need to master beginner data analytics concepts to advanced data visualization and machine learning. 	<ul style="list-style-type: none"> Python Basics for Data Analysis Variables and Data Types Practice Problems
Date TBD TIME TBD (2 hrs)	Service Workbench on AIM-AHEAD Dr. Paul Avillach will provide an overview on how to use the AIM-AHEAD service workbench (SWB) .	<ul style="list-style-type: none"> Supplies Trainees with a simple web application, empowering them to easily deploy and access any cloud workspace from a custom catalog of pre-configured and extendible) environments (R, Jupyter Notebooks, Python, etc...) leveraging all AWS advanced data analysis tools and native security controls. 	<ul style="list-style-type: none"> Have Access to AIM_AHEAD services and products Hands On practices and live sessions
Date TBD TIME TBD (1 hr)	AI/ML/Data Science Module kickoff	<ul style="list-style-type: none"> Learn Foundations of ML/AI Exploring the basic concepts, and industry-standard notations in ML/AI and exploring the real-world contexts for the data science lifecycle. It then progresses to drawing business conclusions from data sets and visualizations. 	<ul style="list-style-type: none"> Trainees will gain hands-on experience with topic like k-means algorithms and apply functions. They will also learn how to predict outcomes using multiple linear regression models, create visual decision trees, and interpret various kinds of ML/AI decision models.
Mid-Late October TIME TBD (2 hrs)	All of Us Data Browser Dr. Brad Malin and team will provide an overview of the All of Us Data Browser discussing access, available data, and best practices	<ul style="list-style-type: none"> Access The All of Us data, research tools, and research projects. The Research Hub has other features for the trainees to explore. 	<ul style="list-style-type: none"> Be able to access and leverage NIH available data. Learn about the other research hub features such as: Researcher Workbench, Workspaces, Notebooks and Dataset Builder.
Date TBD	Machine Learning in	<ul style="list-style-type: none"> Introduces trainees to machine 	<ul style="list-style-type: none"> Explores machine learning

TIME TBD (1 hr)	Healthcare Module kickoff	learning in healthcare, including the nature of clinical data and the use of machine learning for risk stratification, disease progression modeling, precision medicine, diagnosis, subtype discovery, and improving clinical workflows.	methods for clinical and healthcare applications. <ul style="list-style-type: none"> Covers concepts of algorithmic fairness, interpretability, and causality. Discusses application of time-series analysis, graphical models, deep learning, and transfer learning methods to solving problems in healthcare.
Date TBD TIME TBD (1 hr)	Healthcare Applications (Health Equity and Health Disparities) Module kickoff	<ul style="list-style-type: none"> AIM-AHEAD Professional Development Program Introductions 9 Modules to be taken during the training period 	<ul style="list-style-type: none"> Access the AIM_AHEAD and register for the course's modules.
02/24/23 12:00 – 1:00 pm EST (2 hrs)	Trainee Final Group Presentations Trainees will present outcomes from PRIME	<ul style="list-style-type: none"> Report training goals and identify learning achievements 	<ul style="list-style-type: none"> Reflect on outcomes Research outcomes and report if applicable
Graduate: 03/17/22 1:00 pm – 3:00 pm EST Undergrad: 05/19/23 1:00 pm – 3:00 pm EST	Trainee Final Poster Presentations Trainees will present outcomes from PRIME	<ul style="list-style-type: none"> Report training goals and identify learning achievements Reflect on outcomes Research outcomes and report if applicable 	<ul style="list-style-type: none"> Group Poster Presentation <ul style="list-style-type: none"> Individual poster Presentation Selected posters will be presented in the AIM_AHEAD annual conference 2023

ASYNCHRONOUS LEARNING

Provided required asynchronous courses will be available on Canvas. You are required to identify additional courses as electives. These courses must be approved via the **curriculum petition form** (available in files section of PRIME Cohort of 2022 Canvas Page). We have provided a *recommended* timeline to complete the courses before the group presentations. Each item should take approximately 40 hours, for a total of 240 hours of curriculum.

Timeline (10 hrs. curriculum / wk.)	Item	Outcomes	Notes
September 09, 2022 – November 4, 2022	AI for Health Equity Symposium – AIHES2022 Workshop Series Online Workshop Series	<ul style="list-style-type: none"> Learn the importance of AI\ML in health care to address health equity, health disparities and to solve the data, Biases. 	<ul style="list-style-type: none"> Required for the trainees to watch all 12 videos and include a lesson learned in the reports.
	Data Quest Hands-on Practice Online Course	<ul style="list-style-type: none"> Practice: <ul style="list-style-type: none"> Introduction to Python Python Programming For Loops and Conditional Statements in Python Dictionaries, Frequency Tables, and Functions in Python Python Functions and Learn Jupyter 	<ul style="list-style-type: none"> Trainees must sign and finish the free hands-on Python practices and Challenges

		Notebook - Python for Data Science: Intermediate	
November 4, 2022 – January 27, 2023	AI/ML/Data Science Online Course	<ul style="list-style-type: none"> Fundamental topics: Introduction to Machine Learning to Fundamentals of Data Analysis ML/AI Techniques: Clustering and Principal Component Analysis to Support Vector Machines And Advance topics from Natural Language Processing to Deep Neural Networks 	<ul style="list-style-type: none"> Basic topics available for undergraduate students Advanced topics available for graduate students
	Machine Learning in Healthcare Online Course	<ul style="list-style-type: none"> Trainees will learn about 21 topics including: Overview of Clinical Care, Physiological Time-Series, Clinical NLP, Survival modeling, Fairness, Interpretability and more. 	<ul style="list-style-type: none"> Trainees will sign up for a canvas course that is going to be self-thoughts. Evaluation forms will determine the amounts invested in learning about Machine learning in health care.
	Healthcare Applications for Health Equity / Health Disparities Online Course (AAPDP)	<ul style="list-style-type: none"> AIM-AHEAD Professional Development Program Introductions 9 Modules to be taken during the training period. 	<ul style="list-style-type: none"> This course will have recommended assignments and quizzes assigned to the trainees after each course module.
January 27, 2023 – end of program <small>(Grad – March 8, 2022; Undergrad – May 8, 2023)</small>	Electives Approved via curriculum petition template	<ul style="list-style-type: none"> Elective courses that is aligned with the AIM_AHEAD objectives. 	<ul style="list-style-type: none"> Must be approved via Curriculum Petition Form

TIME BREAKDOWN

There is an expectation that trainees by accepting this award are committing to the stated weekly effort allocations (20 hours / week: Graduate level and higher; 15 hours / week: Undergraduate-level), but the hours for each type of activity is flexible to the trainee. You must complete a weekly tracker and attestation that includes a detailed reflection of your takeaways from that week's training activities. This is to be signed by yourself weekly and signed by your mentor monthly. The signed form must be uploaded through the assignments tab on the PRIME Cohort of 2022 Canvas Page. **Total contracted hourly commitment: 480 hours.**

Item	Total hrs.
AI for Health Equity Symposium – AIHES2022 Workshop Series	40
Data Quest Hands-on Practice	40
AI/ML/Data Science	160
Machine Learning in Healthcare	40
Healthcare Applications	40
Electives	40

Reporting & Reflecting	64
Final Presentation	40
Synchronous meetings	16

We recognize that the time spent on different activities will fluctuate from week to week. However, we have provided an *example* of a rough breakdown of your **weekly effort**.

GRADUATE-LEVEL

Graduate-level trainees will focus relatively evenly on training curriculum, interacting with Mentors and Mentees, and working on project-based activities during their **6-month program**. The remaining time should be spent connecting with institutional and peer-mentees, as well as reflecting and compiling progress reports. Project-based activities will be largely defined by their institutional mentor and dependent on the project for which the trainee is supporting. Reflection and reporting category includes completing weekly effort tracker.

Example

Hours/week	Activities
10	Curriculum / courses
4	Mentors / mentee interactions
4	Project-based activities (if applicable)
2	Reflection & reporting
20	Total

UNDERGRADUATE-LEVEL TRAINEES

Undergraduate-level trainees will focus primarily on training curriculum during their **8-month program**. The remaining time should be spent connecting with institutional and peer-mentors as well as reflecting and compiling progress reports. Reflection and reporting category includes completing weekly effort tracker.

Example

Hours/week	Activities
10	Curriculum / courses
3	Training activities
	AIM-AHEAD activities
1	Mentor / mentee interactions
1	Reflection & reporting
15	Total

FINAL PRESENTATIONS

You will give two presentations at the end of your PRIME program. One presentation will be a group presentation and one will be an individual poster.

GROUP PRESENTATIONS

All students will be placed into groups of roughly 5-6 students. These groups will each be assigned one of the asynchronous courses/modules to cover.

Your presentation should include

- Takeaways from the course/module
- Challenges
- Use cases / applications
- Next steps

INDIVIDUAL POSTER PRESENTATIONS

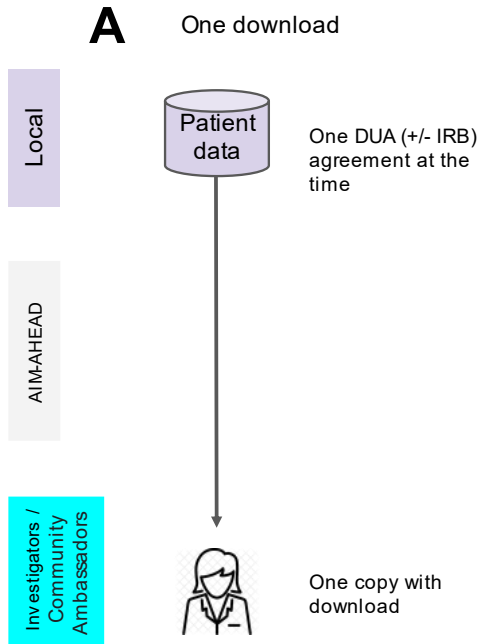
- You will present a poster of your design on either (1) Training applications on personal research project, or (2) Training outcomes (methods). Graduate students are expected to present on (1) while undergraduate students will likely present on (2) with the option of (1).
- *Optional:* Graduate students presenting on training applications on personal research project are encouraged to include a PRIME undergraduate student as a co-author to your poster in addition to your mentor.
- You will be provided a template and a few examples. **Your poster should include Title, Author(s), Abstract, Methods, Results, Conclusion/Discussion.** The template will include a mandatory funding disclosure; if the information in the poster was supported by other external funding, you must also include the associated funding disclosure in addition to the AIM-AHEAD disclosure in the template.
- All figures and tables should include a caption. Graphs must include a legend.
- Graduate students will present posters at the end of their 6-month program and undergraduate students will present posters at the end of their 8-month program.
- Your presentation will be recorded.

EXPERIENTIAL LEARNING ASSESSMENTS & FEEDBACK

Experiential feedback-based learning promotes critical thinking and problem-solving skills in learners with different career trajectories. PRIME program employs assessments that **allow students to develop more than just managerial and soft skills**. The assessments allow students to explore ethical and personal values, such as social responsibility and commitment, which impact their personal lives as much as their professional outlook.

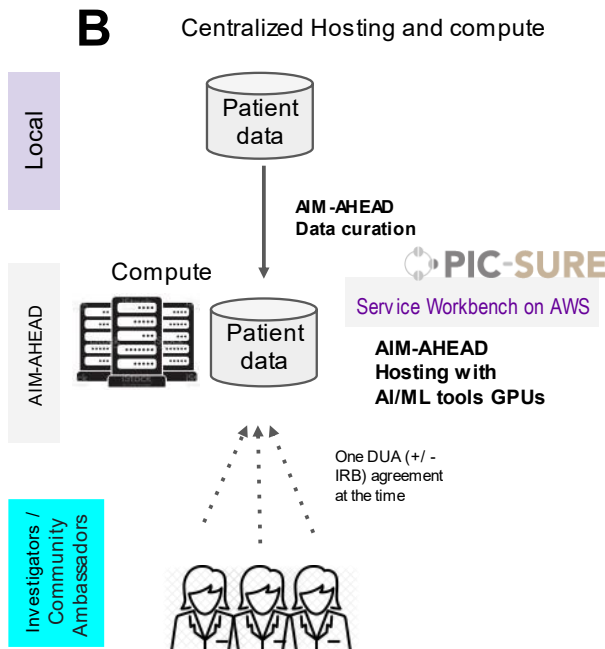
DATA ACCESS & USE

5 MODELS:



AIM-AHEAD offering

Front door to get started (concierge services)
 Listing of existing datasets
 Assistance to IRB and DUAs



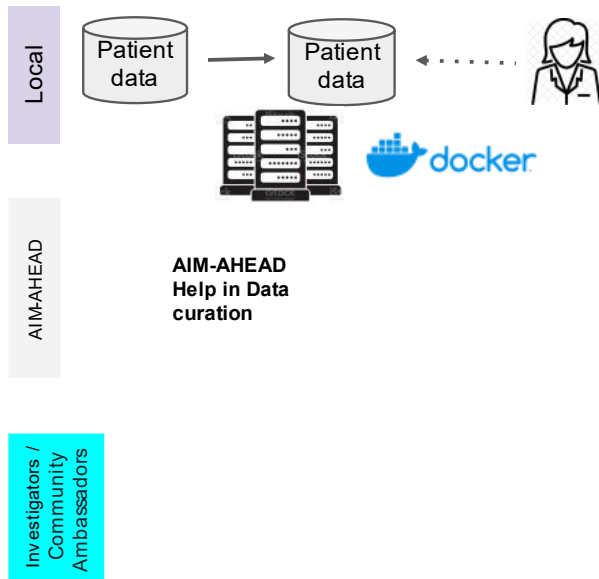
AIM-AHEAD offering

AIM-AHEAD Research Fellowship

Platform for multiple FOA:
 Data Science Training Core

Hosting all OCHIN data?
 Data and Research Core

C Local data analysis by local investigators



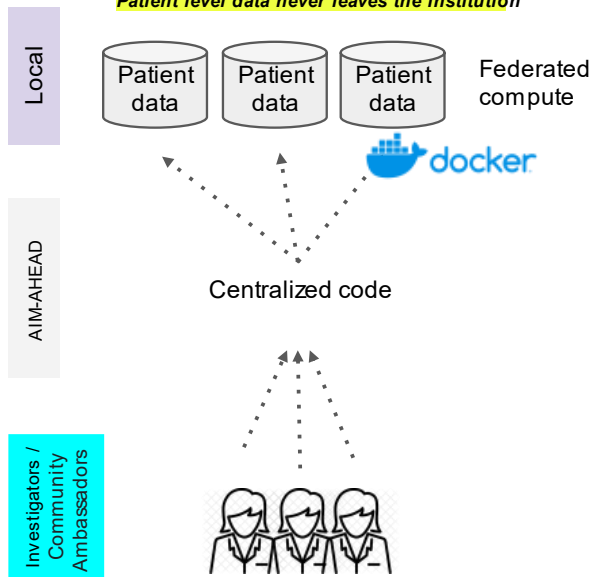
AIM-AHEAD offering

Concierge services to **Help in Data curation for local investigators accessing data** curation

Concierge services to **assist with and troubleshoot analysis**

D Creating an AIM-AHEAD Federated Networks

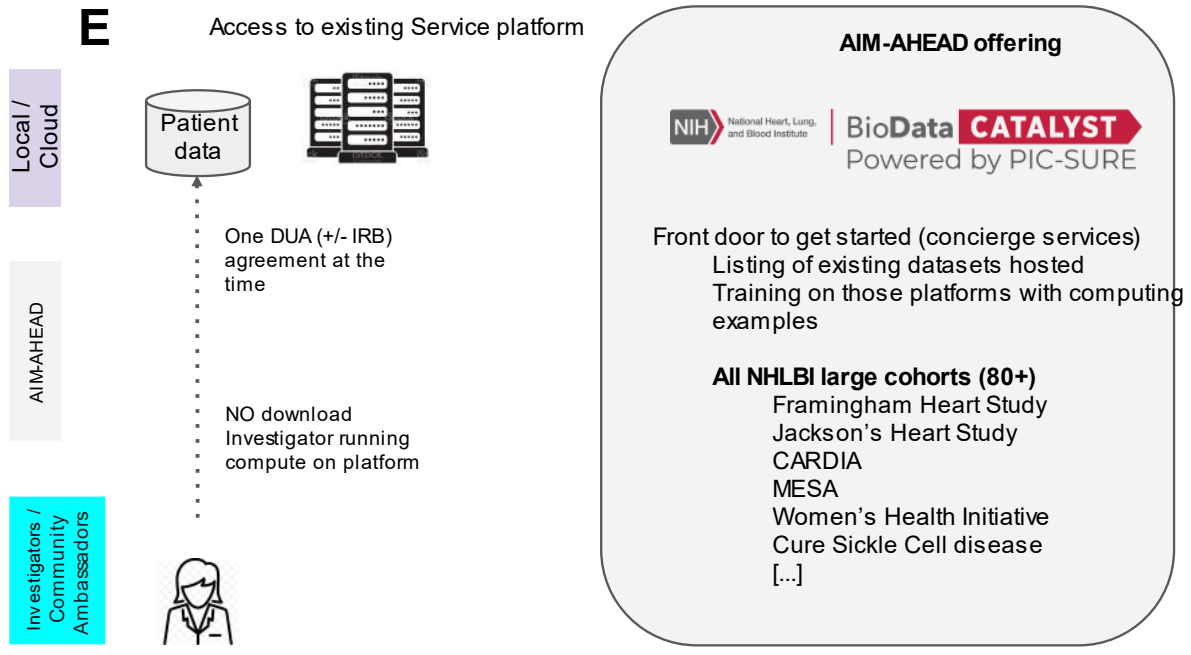
Patient level data never leaves the institution



AIM-AHEAD offering

Ongoing discussions with Speero
Will contact all other HUBs

Concierge Services to **assist with data curation, analysis and regulatory considerations**



[AWS SERVICE WORKBENCH \(AWS SWB\)](#)

[Service Workbench on AWS](#) (SWB) promotes infrastructure equity by providing the same analytical tools and level of access to researchers and students. The only infrastructure required to fully leverage this resource is a simple laptop and an internet connection.

SWB provides researchers and students with a user-friendly environment to configure and deploy their own secure cloud-computing environment in a few clicks. SWB is the first [Apache 2.0 open-source native cloud computing platform](#) that provides a modular and scalable solution to the supply of computing environments for researchers and students. The platform supplies students and research teams with a simple web application, empowering them to easily deploy and access any cloud workspace from a custom catalog of pre-configured and extendible) environments (R, Jupyter Notebooks, Python, etc...) leveraging all AWS advanced data analysis tools and native security controls.

To get started with an analysis, end-users are only required to connect to the web application and select their desired configuration. The research workspace will be deployed in two clicks, selecting first the type of workspace and second the most applicable configuration for their analysis in terms of instance type, memory, CPUs, and GPUs. Researchers will have access to the computing power they need, regardless of the technical underlying complexity of it. Moreover, there is a growing open community supporting various SWB workspaces, which enables the deployment of any type of computing workspaces.

Examples of the resources available - out-of-the-box - in SWB: [AWS SageMaker](#) instances that work with widely used Jupyter Notebook formats. Moreover, SageMaker instances come with staple ML/DL libraries (e.g., TensorFlow, PyTorch, MxNet), allowing savvy users to get started right away. On the opposite, non AI/ML experts can discover all [pre-configured computing environments featuring many tutorials and analysis examples using public data in the form of Jupyter Notebooks](#) for anyone to start learning using those resources. This AI/ML service accelerates innovation with purpose-built tools for every step of ML development, including labeling, data

preparation, feature engineering, statistical bias detection, auto-ML, training, tuning, hosting, explainability, monitoring, and workflows. SageMaker helps data scientists and developers to prepare, build, train, and deploy high-quality ML models quickly by bringing together a broad set of capabilities purpose-built for ML, and enables the Fellows to develop and serve their own models.

AIM-AHEAD Service Workbench User Guide: <https://docs.google.com/document/d/1nrpLLpmm66-G7Mo-BOBUkGu-DN7fD8YCK3SL9CrPhp0/edit?usp=sharing>

AIM-AHEAD Connect Service Workbench discussion group is here: <https://connect.aim-ahead.net/group/public/2eB9558>

Questions regarding AIM-AHEAD Service Workbench may be directed to: https://helpdesk.aim-ahead.net/ticket/create/Service_Workbench