

Introduction to the *All of Us* Researcher Program

INSTRUCTORS: **Paul Harris, Ph.D., FACMI, FIAHSI**

Professor, Biomedical Informatics, Biomedical Engineering, & Biostatistics
Director, Office of Research Informatics
Vanderbilt University Medical Center
paul.a.harris@vanderbilt.edu

Qingxia ‘Cindy’ Chen, Ph.D.

Vice Chair of Education, Department of Biostatistics
Professor, Biostatistics, Biomedical Informatics, and Ophthalmology & Visual Sciences
Vanderbilt University Medical Center
cindy.chen@vumc.org

Other members from the *All of Us* Data and Research Center

SCHEDULE: Monday, July 22 – Friday, July 26, 2024
1:00 PM – 4:00 PM (Central Daylight Time)

OVERVIEW

The *All of Us* Research Program is a large-scale initiative that collects and studies data from over one million participants in the United States. This program gathers diverse data types, such as survey data, electronic health records, physical measurements, mobile health data (Fitbit), whole genome sequences, and array data. As a freely accessible resource, it plays a crucial role in improving health and advancing precision medicine. Researchers can access and analyze the *All of Us* data via the Researcher Workbench, a secure, Google cloud-based platform. This short course will cover (1) an overview of the Researcher Workbench and its built-in tools, (2) the process of identifying and extracting variables, and (3) case studies illustrating the analysis of various data types using the Workbench.

Upon completion of this course, students should be able to:

- Understand the goals and scope of the *All of Us* Research Program and its significance in advancing precision medicine.
- Navigate the Researcher Workbench confidently and effectively utilize its built-in tools for data analysis.
- Identify and extract relevant variables from diverse data types within the *All of Us* dataset.
- Comprehend some featured workspace studies that utilize a range of data types accessible in the Researcher Workbench, including survey data, electronic health records, physical measurements, mobile health data (Fitbit), whole genome sequences, and array data.
- Interpret the results of case studies to gain insights into potential applications of the *All of Us* data in health research.

- Collaborate with fellow researchers to design and execute research projects using the *All of Us* dataset.
- Understand and adhere to data security and privacy guidelines while working with sensitive information from the *All of Us* Research Program.

PREREQUISITES

1. Basic programming knowledge in R, Python, or RStudio.
2. Set up an *All of Us* research account and complete the required training. Students **must** have an *All of Us* researcher workbench account set up by the start of the course. To obtain an account, complete all the steps at <https://www.researchallofus.org/register/>. Because of the multiple requirements to be met, including participation in an hour-long training session, we strongly advise initiating your account set-up by **June 14** at the latest.
3. Creation and duplication of workspaces in advance of the course. Instructions will be shared closer to the beginning of the event.

Additional note: The *All of Us* Research Program is currently available only to individuals at academic, nonprofit, or health care institutions. Before registering for the course, please ensure your institution has an agreement on file, which can be found at <https://www.researchallofus.org/institutional-agreements/>. If your organization does not have an agreement, you can initiate one at <https://redcap.pmi-ops.org/surveys/?s=7N7TA9AHAA>. If you encounter technical issues during registration, please contact support@researchallofus.org.

SOFTWARE

The three software programs listed below will be utilized throughout the course. These programs are accessible within the Workbench for all registered researchers. Gaining a fundamental understanding of these software programs will enhance your comprehension of the case studies covered in this course.

- R: <https://cran.r-project.org/>
<http://cran.r-project.org/doc/manuals/r-release/R-intro.html>
- Python: <https://www.python.org/>
- Jupyter Notebook: <https://jupyter.org/>
- [RStudio](#)

SUPPLEMENTAL MATERIALS

Case studies for this course, also known as featured workspaces, can be found on the *All of Us* Researcher Workbench. To view a publicly available table of contents of these materials, visit <https://support.researchallofus.org/hc/en-us/articles/360059633052-Featured-Workspaces-Table-of-Contents>.

COURSE OUTLINE

Day 1	<p><u>Topic</u></p> <ul style="list-style-type: none"> ● Introduction to the course and the <i>All of Us</i> Research Program ● Overview of data and resources ● How to collaborate with other researchers on the Researcher Workbench and set up your workspace ● Constructing a dataset using analytic tools ● Overview of persistent disk, storage bucket, and cost ● Publishing results derived from the <i>All of Us</i> Research Program ● Hands-on lab exercises 	<p><u>Instructor(s)</u> Qingxia ‘Cindy’ Chen</p> <p><u>Teaching Assistants</u> Shawn Garbett Andrew Guide Sarah Feng</p>
Day 2	<p><u>Topic</u></p> <ul style="list-style-type: none"> ● Case Study 1: Surveys ● Understanding and working with survey data in the <i>All of Us</i> Research Program ● Introduction to OMOP tables and data structure for Survey Data ● Introduction to SQL ● Hands-on lab exercises 	<p><u>Instructor(s)</u> Qingxia ‘Cindy’ Chen Shawn Garbett</p> <p><u>Teaching Assistants</u> Andrew Guide Sarah Feng</p>
Day 3	<p><u>Topic</u></p> <ul style="list-style-type: none"> ● Case Study 2: Electronic Health Records (EHRs) ● Understanding and working with EHR data in the <i>All of Us</i> Research Program ● OMOP tables and data structure for EHRs ● Hands-on lab exercises 	<p><u>Instructor(s)</u> Cathy Shyr Lina Sulieman</p> <p><u>Teaching Assistants</u> Shawn Garbett Andrew Guide Sarah Feng</p>
Day 4	<p><u>Topic</u></p> <ul style="list-style-type: none"> ● Case Study 3: Genetic Study ● Introduction to genetic data in the <i>All of Us</i> Research Program ● Hands-on lab exercises 	<p><u>Instructor(s)</u> Genevieve Brandt</p> <p><u>Teaching Assistants</u> Michael Lyons Shawn Garbett Andrew Guide Sarah Feng</p>
Day 5	<p><u>Topic</u></p> <ul style="list-style-type: none"> ● Case Study 4: Mobile Health Study (Fitbit) ● Introduction to Fitbit data in the <i>All of Us</i> Research Program 	<p><u>Instructor(s)</u> Hiral Master</p>

	<ul style="list-style-type: none">• Hands-on lab exercises• Course wrap-up and Q&A	Aymone Kouame <u>Teaching Assistants</u> Hunter Hollis Shawn Garbett Andrew Guide Sarah Feng
--	---	--