



The *All of Us* Research Program & the Researcher Workbench: an Invitation to Researchers



Romuladus E. Azuine, Dr.PH., M.P.H., RN
Branch Chief, Researcher Platforms Branch
& Director, AoU Data and Research Center
*4th National Big Data Health Science Conference,
University of South Carolina, February 11, 2023*



OVERVIEW

At the end of my presentation, you will be able to:

- What is the *All of Us* Research Program?
- What data are participants contributing?
- What data are available to researchers now?
- How do I get access to the data?
- What tools are available for analyzing the data?
- Who else is using this data, and how?

- **Partners**
 - 100+ partner organizations
 - 650+ sites collecting samples & measurements

All of Us Community and Provider Partner Network (as of August 2022)



Note: These are not approved lockups and should not be repurposed on assets.

All of Us Consortium Members (as of August 2022)

The Participant Center



Communications & Engagement

WONDROS



HPO Network (Health Care Provider Organizations)

RMCs

All of Us California



Illinois Precision Medicine Consortium



All of Us New England



Trans America Consortium



New York City Consortium



All of Us Southern Network



All of Us Southeast Enrollment Center



HPO Lite



All of Us Wisconsin



All of Us Pennsylvania



University of Arizona and Banner Health



FQHCs (Federally Qualified Health Centers)



VA Medical Centers



Participant Technology Systems Center (PTSC)



Biobank



Data & Research Center (DRC)



Genomics Partners



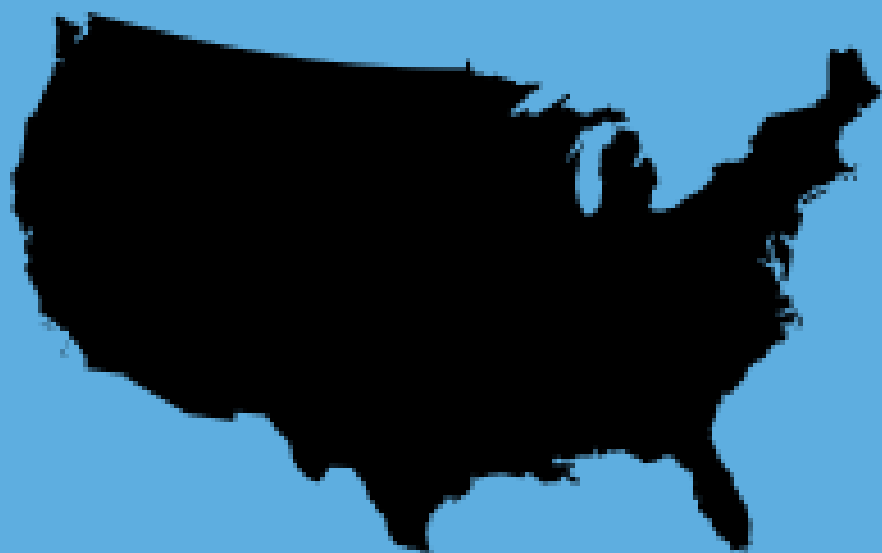
What is the *All of Us* Research Program?

[ResearchAllofUs.org]

One of the largest, richest, most diverse biomedical datasets of its kind

Inviting

1 Million or more people
across the
United States



Enrolled **596,000+** participants

80%
are from
communities
underrepresented in
biomedical research



50%
are from racial
and ethnic
minority groups

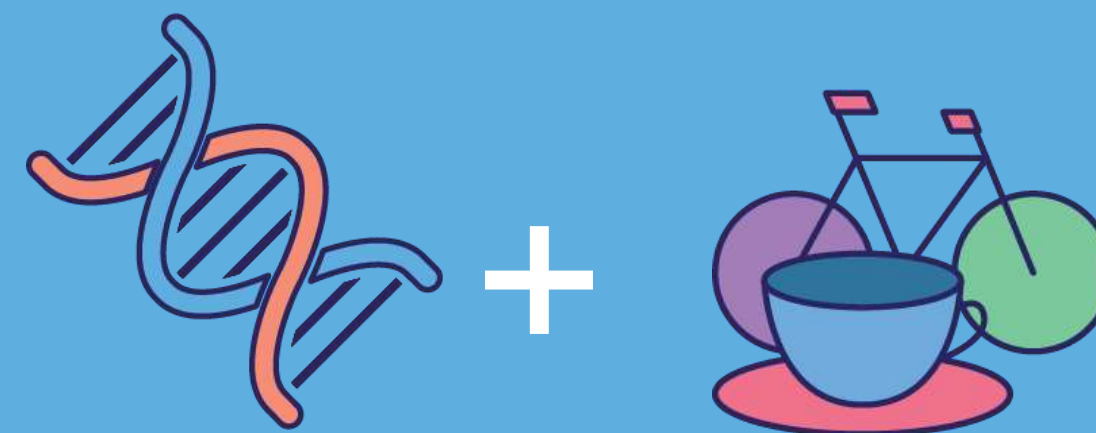
Data as of January 2023

Enables research discoveries that drive more precise approaches to care

Engages **people & communities who have been left out of medical research** in the past



Combines **biological factors and social determinants** on a large, inclusive scale



Easily accessible to any researcher with a secure internet connection and data use agreement



Follows participants as they move, age, and grow



By securely capturing a range of different data types



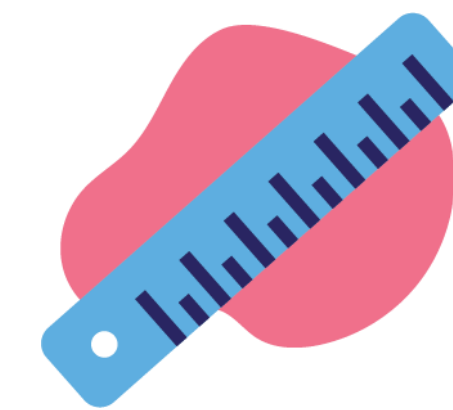
Surveys

(from 372,000+ participants)
Including lifestyle, access to care, medical history, and **data from nearly 100,000 participants on their experiences during the COVID-19 Pandemic**



Electronic Health Records

(from 349,000+ participants)
Standardized to OMOP common data model



Physical Measurements

(from 306,000+ participants)
Blood pressure, heart rate, BMI and more



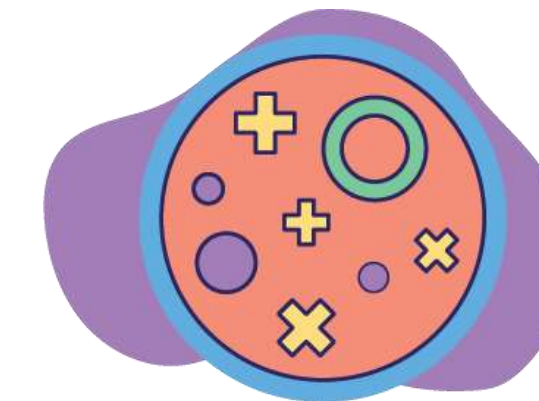
Wearables

(from 12,800+ participants)
Physical activity and heart rate from participants who have connected their Fitbit devices



Genomics


(Nearly 100,000 whole genome sequences) and 165,200 shorter DNA samples known as genotyping arrays
Only available via the Controlled Tier



Biosamples

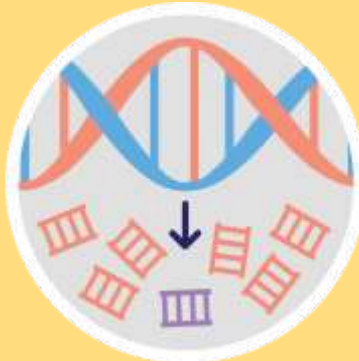
Over 400, 000

Including the first batch of genomic data




Genomics


Only available via
the Controlled Tier



98,600
Whole Genome
Sequences




165,200
Genotyping
Arrays



593,500,000+
Unique Variants








Represented and searchable
through public genomic
data browser








**Genomics
Analysis Tools**

Hail and PLINK
in addition to R, Python,
and Jupyter Notebooks




Genomic Data is Paired with Rich Phenotypic Data






81,000+
Have Whole Genome Sequences + Electronic Health Records
+ Physical Measurements + Survey Responses



98,000+
Have Whole Genome Sequences + Physical Measurements +
Survey Responses



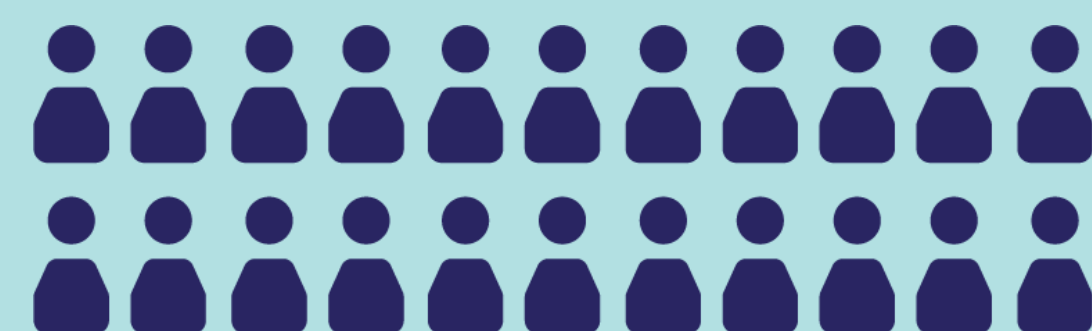
81,000+
Have Whole Genome Sequences + Electronic Health Records



3,300+
Have Whole Genome Sequences + Fitbit Records
Representing >25% of all participant Fitbit records

And data from the recurring COVID-19 Participant Experience (COPE) survey

COPE Survey Data Currently Available in the Researcher Workbench



More than **100,300** participants responded to



one or more of the **six** COPE surveys administered between May 2020 and March 2021

*45,547 participants responded to three or more surveys.

Topics Covered



Mental Health

The COPE data represents the biggest infusion of mental health data into the Researcher Workbench so far. Insights include:



62%
felt bothered by sleeping problems



53%
felt nervous and anxious



94%
have someone to love and make them feel wanted



95%
have someone to have a good time with

Embedded in the survey were resources for participants struggling with emotional distress, including suicide prevention tools.

*Participants could select multiple symptoms.

And making them accessible to researchers across stages and settings

Our Researchers



3,930+

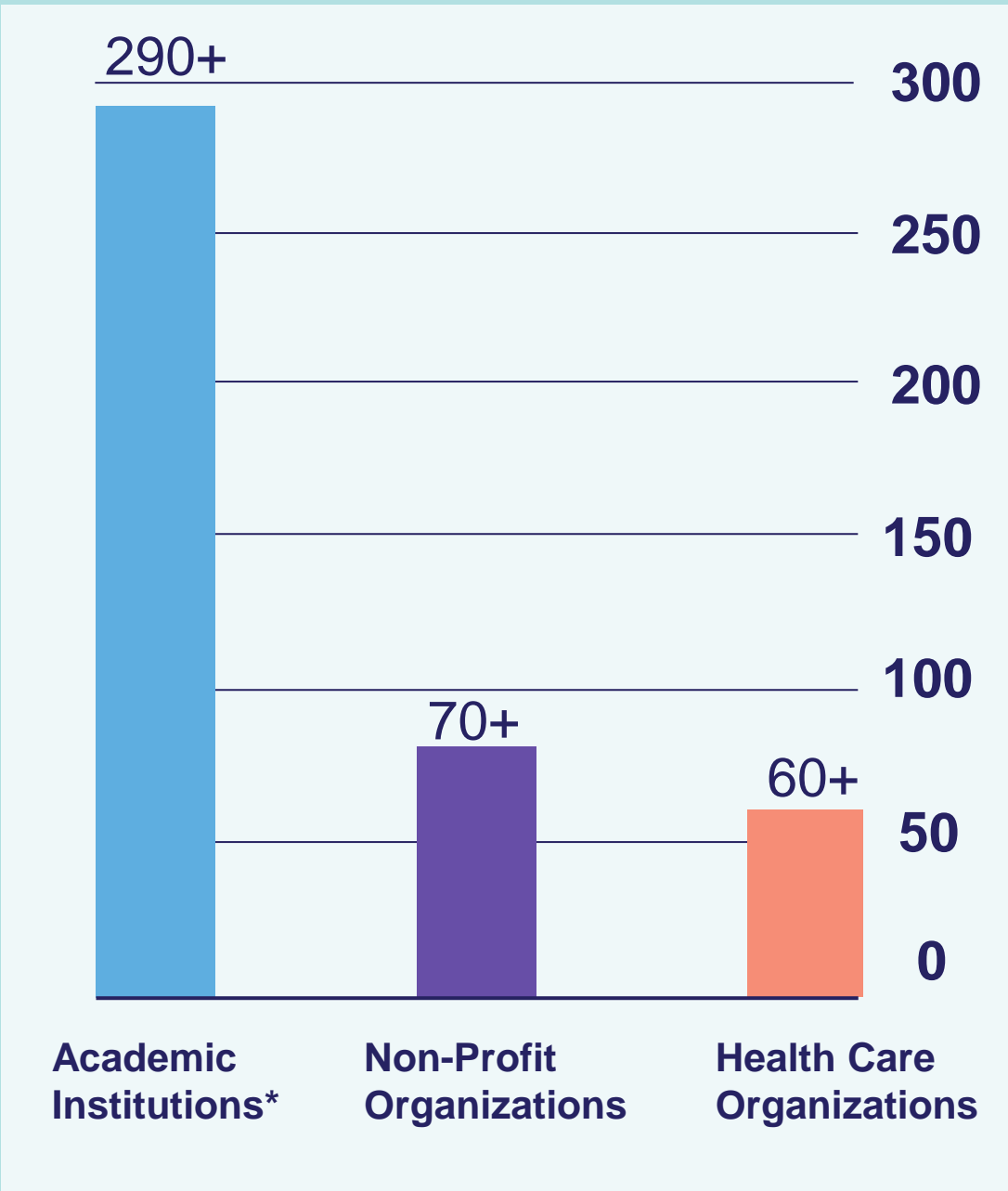
**Registered
Researchers**

across a range of
institutional roles
and career stages



450+

Institutions



*Includes 60+ Historically Black Colleges and Universities and Hispanic-Serving Institutions.

Figures accurate as of September 2022

Figures accurate as of January 2023

Research Currently Underway



3,310+

Active projects



115+

**Publications in
peer-reviewed
journals**



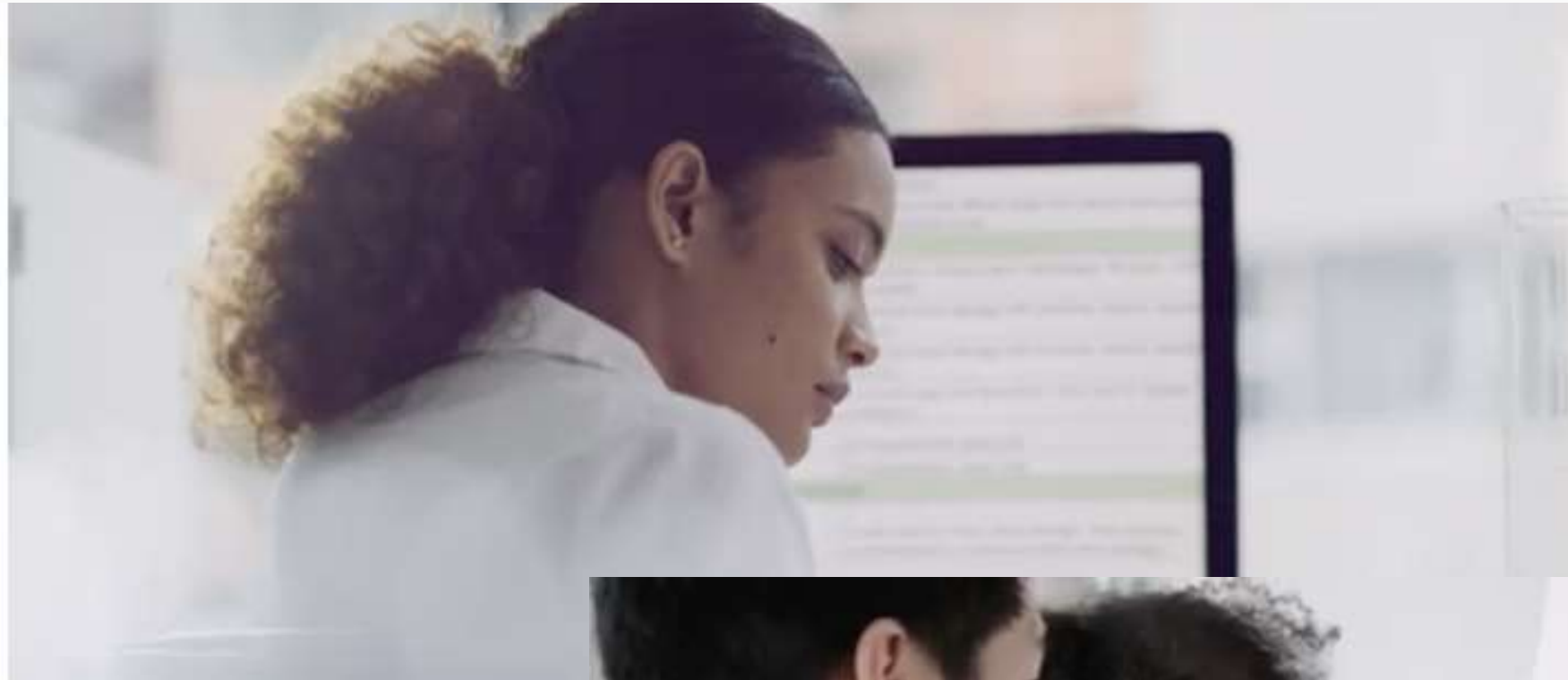
Top conditions being studied

In the Researcher
Workbench include:

- Cardiovascular disease
- Hypertension
- Mental Health
- Cancer
- Diabetes

Figures accurate as of January 2023

Building a diverse researcher cohort



- **Creating a demographically diverse researcher cohort** that promotes responsible and ethical use of data, returns value to participant communities, and accelerates research impact.
- **Encouraging student assemblies and early-stage investigators** to bring fresh, creative perspectives & innovative research outcomes.
- **Ensuring access for researchers from various institutions/organizations** to establish a truly equitable resource for all.

With aggregated overviews and interactive previews available to everyone

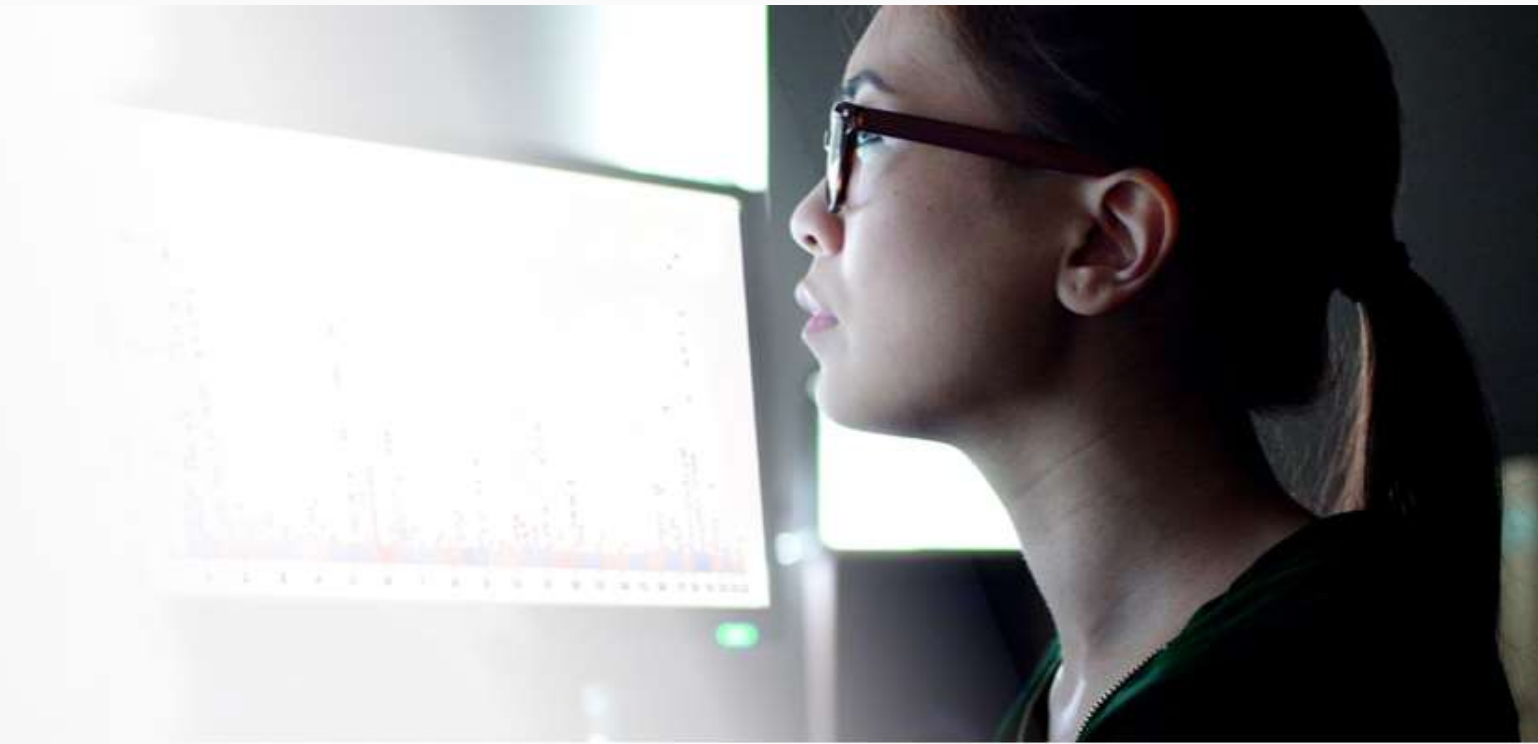
ResearchAllofUs.org

Welcome to the *All of Us* Research Hub

The *All of Us* Research Program, led by the National Institutes of Health, is building one of the largest biomedical data resources of its kind. The *All of Us* Research Hub stores health data from a diverse group of participants from across the United States.

Registered researchers can access *All of Us* data and tools to conduct studies to help improve our understanding of human health.

REGISTER FOR ACCESS



Search Across Data Types

Keyword Search

Data includes 372,380 participants and is current as of 6/6/2022.



EHR Domains

<div>Conditions</div> <div>24,315</div> <div>medical concepts</div> <div>227,740 participants in this domain</div> <div>View Conditions</div>	<div>Drug Exposures</div> <div>29,166</div> <div>medical concepts</div> <div>214,040 participants in this domain</div> <div>View Drug Exposures</div>	<div>Labs & Measurements</div> <div>15,309</div> <div>medical concepts</div> <div>227,280 participants in this domain</div> <div>View Labs & Measurements</div>	<div>Procedures</div> <div>29,176</div> <div>medical concepts</div> <div>221,860 participants in this domain</div> <div>View Procedures</div>
---	---	---	---

Data Snapshots

Data Snapshots showcase the scale and diversity of the *All of Us* Research Program participant cohort. The snapshots provide participant demographics, geographic distribution, and more. We update the snapshots daily.



580,000+

Participants



341,000+

Electronic Health Records



421,000+

Biosamples Received

VIEW MORE DATA SNAPSHOTS

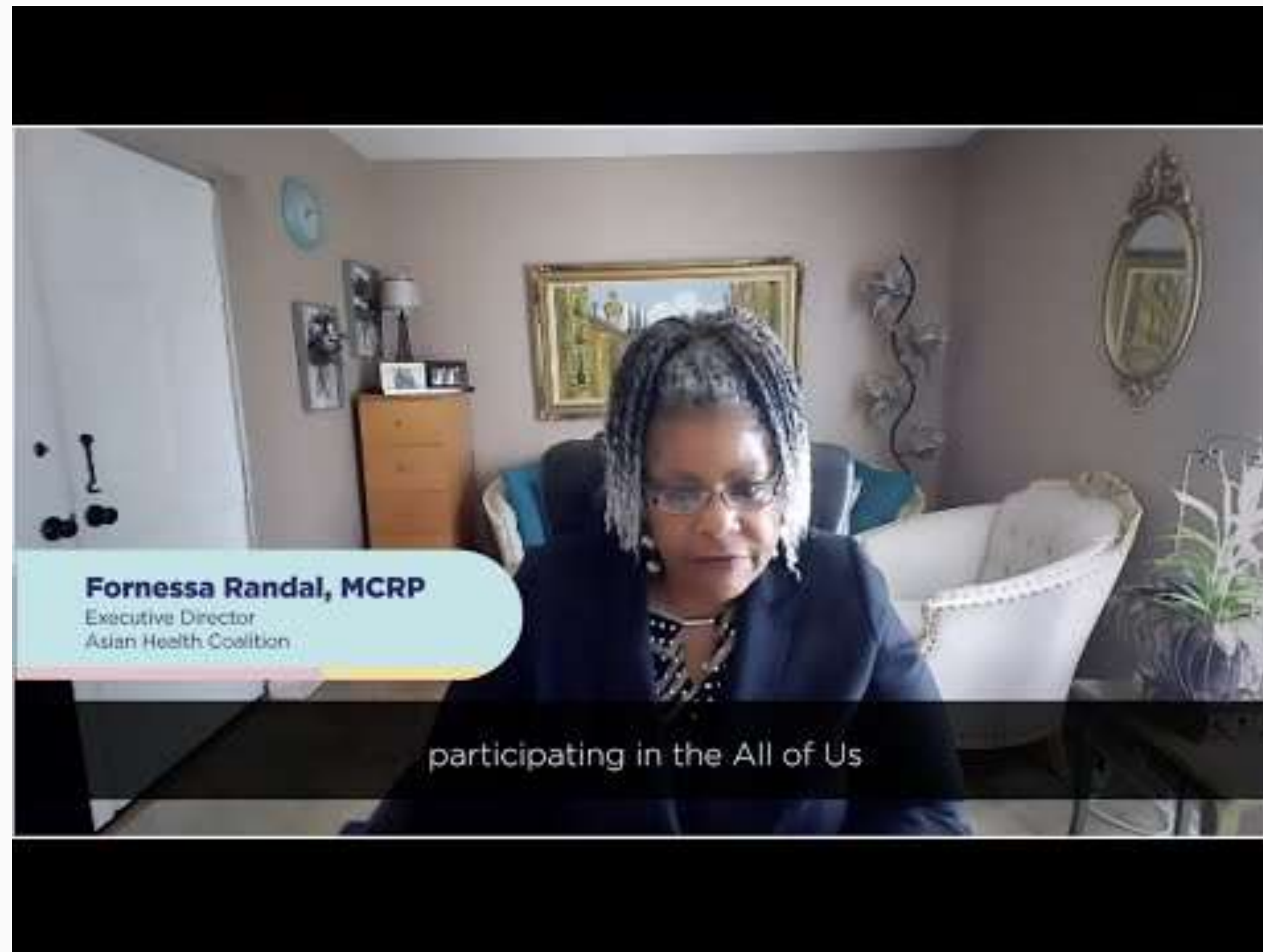
Genomics

<div>Genomic Variants</div> <div>98,600</div> <div>participants in the Whole Genome Sequencing (WGS) dataset</div> <div>165,140</div> <div>participants in the Genotyping Array dataset</div> <div>View Genomic Variants</div>
--

Physical Measurements and Wearables

<div>Physical Measurements</div> <div>8</div> <div>Physical Measurements</div> <div>311,300 participants in this domain</div> <div>Participants have the option to provide a standard set of physical measurements.</div> <div>View Physical Measurements</div>	<div>Fitbit</div> <div>4</div> <div>Fitbit Measurements</div> <div>12,880 participants in this domain</div> <div>Fitbit data includes heart rate and activity summaries.</div> <div>View Fitbit</div>
---	---

Conducting research with the *All of Us* Research Workbench



Data Access and Analytical Tools

Tiered access levels enable discovery



Available to
anyone

Public Tier

Anyone can visit ResearchAllofUs.org (the *All of Us* Research Hub) to learn more about the data available for research and explore aggregated participant data and summary statistics, with identifiers removed. Public resources include:

- **Data Snapshots:** Aggregated, public-facing overviews of participant characteristics and data types
- **Data Browser:** Interactive preview into the larger *All of Us* dataset through public-facing aggregate data
 - Currently includes participant-provided information such as surveys and physical measurements taken at the time of participant enrollment as well as electronic health record (EHR) data
- **Survey Explorer:** Collection of participant survey questions and source information
- **Research Projects Directory:** Information about each research project within the Researcher Workbench

RESEARCHER WORKBENCH



Available to
registered
researchers

Registered Tier

Registered users can access curated, individual-level data and a variety of research tools to conduct a wide range of studies **and NEW data** including:

- Initial Social Determinants of Health (SDOH) data
- Additional COVID-19 EHR data



Surveys



Electronic
Health Records



Physical
Measurements



Wearables

Data have been processed to protect participant privacy

Controlled Tier

Registered users with amended institutional agreements can access **all of the data in the registered tier plus**

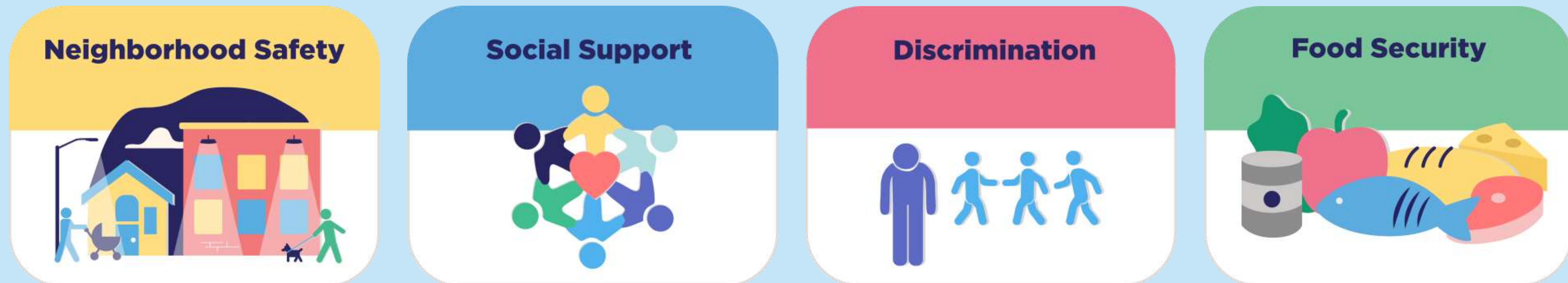
- Expanded demographics
- Unshifted event dates
- Genomic data derived from WGS (~100k participants) and array genotyping (165,200 participants)
- CRAM files which complement the genomic variant calls and enable capabilities such as custom variant evaluation and visualization
- intensity files (iDat) for custom array analyses beyond variants

Linked to drive new insights

By readily connecting biological and social determinants of health data on a large, inclusive scale and following participants as they move, age, and grow, the *All of Us* dataset is driving new insights into health and disease.

Initial Social Determinants of Health (SDOH) Survey Data

Including:



From 57,600+ responses to SDOH surveys

Figure accurate as of June 2022

And this is just the beginning

In the coming years, *All of Us* expects to enroll one (1) million participants and continue adding new data to the dataset, including:



**More Wearables
Data**



**More
Surveys**

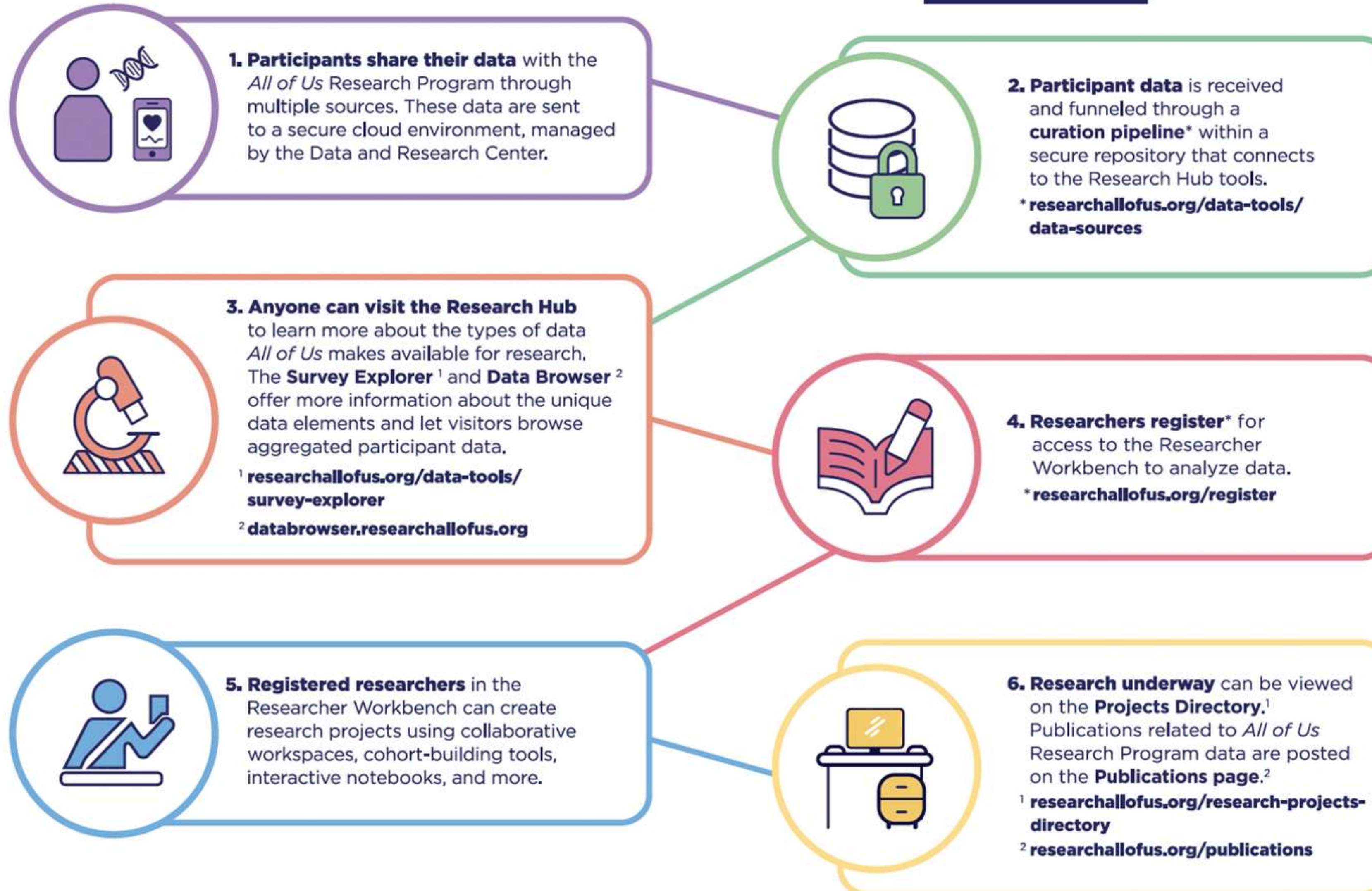


**New Linkages to
External Data
Sources**



**Ancillary
Study Data**

All of Us Research: How It Works



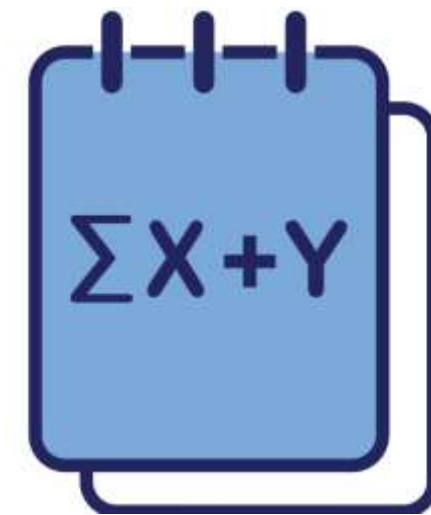
Powerful analytical tools & user support



WORKSPACES

USES:

- Organizing research projects
- Collaboration
- Transparency



NOTEBOOKS

USES:

- Analyses
- Queries

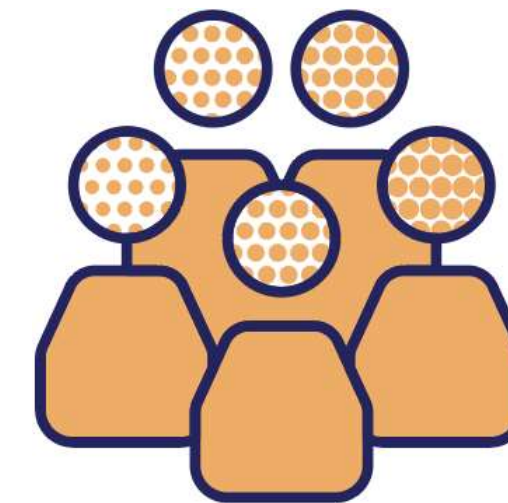
Uses programming languages R and Python



DATASET BUILDER

USES:

- Pre-populated analyses
- Dataset previews



COHORT BUILDER

USES:

- Cohort creation



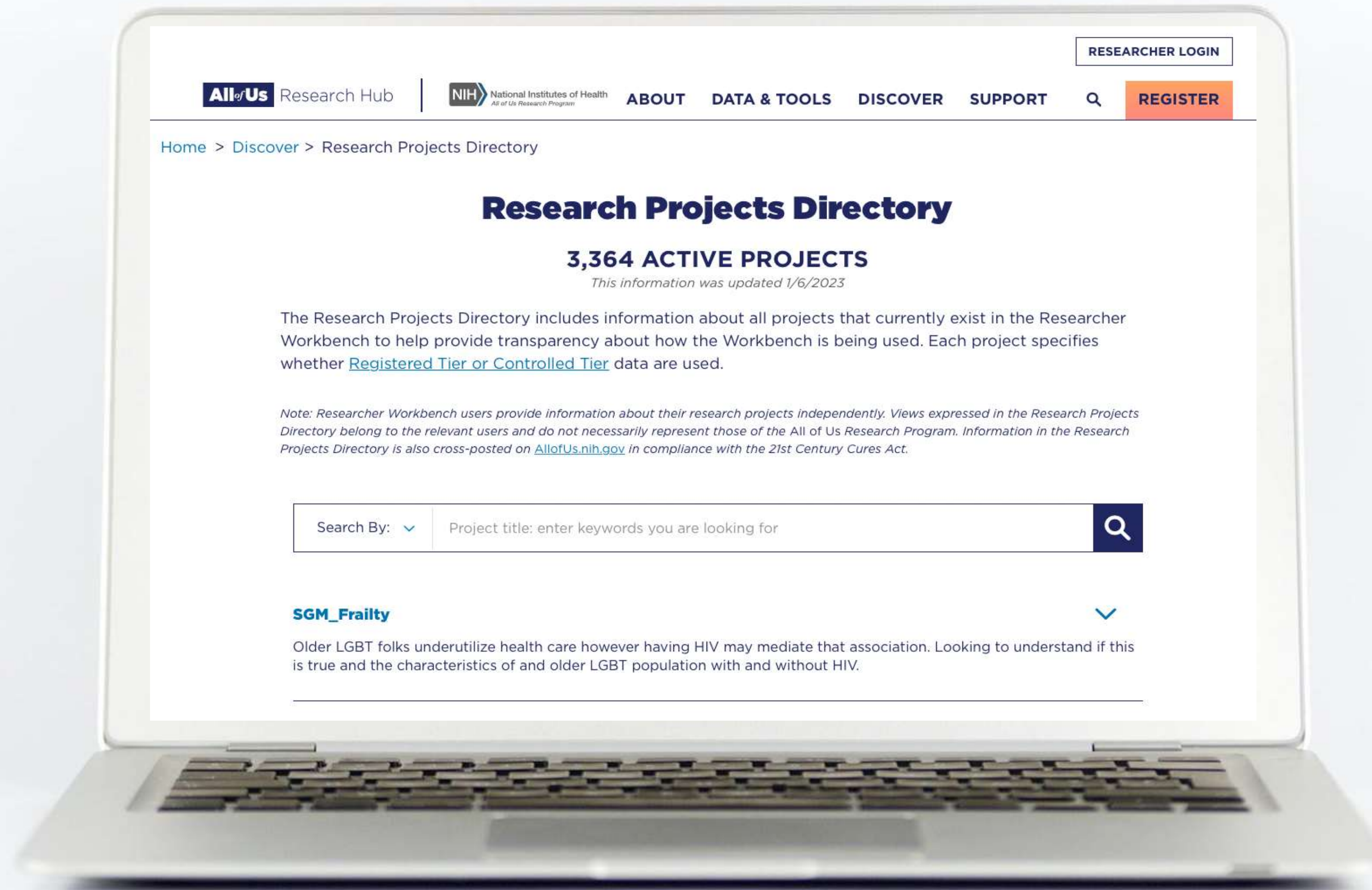
USER SUPPORT HUB

USES:

- Workbench training and resources
- Articles on Data and Tools
- Live office hours

Projects

Explore how other researchers are using the data



researchallofus.org/research-projects-directory

Testimonials from registered researchers

Predictive analytics for glaucoma



Sally Baxter, MD, MSc
Assistant Professor, Shiley Eye Institute
University of California San Diego Health

How diverse data can power more impactful research



Jason Karnes, PharmD, PhD, BCPS, FAHA
Director of Scientific Programs
University of Arizona/Banner Health

Studying health care access and utilization among adult cancer survivors



Jie Chen, PhD
Professor and Division Chief,
Biostatistics and Data Science
Medical College of Georgia, Augusta University

Testimonials from researcher ambassadors

Using the Researcher Workbench to aid discoveries in women's health and health disparities research



Ky'Era Actkins, PhD
Postdoctoral Researcher, Vanderbilt University
All of Us Researcher Ambassador

How the *All of Us* dataset and tools can empower early career researchers and advance research in communities underrepresented in biomedical research



Isaias Torres
Student, College of Osteopathic Medicine, Rocky Vista University
All of Us Researcher Ambassador

Testimonials from researcher ambassadors, cont'd

Understanding the behavioral health needs of people with HIV



Josh Maticotta, Psy.D., CAHIMS
Assistant Dean of Assessment & Strategic Initiatives
College of Health Sciences, Western University of Health Sciences
All of Us Researcher Ambassador

Meaningful inclusion of LGBTQ+ people in research



Scott Emory Moore, PhD, MSN, APRN, AGPCNP-BC, FAAN
Assistant Professor, Frances Payne Bolton School of Nursing
Case Western Reserve University
All of Us Researcher Ambassador

Create an *All of Us* account today

Register to be an *All of Us* Researcher

STEP
1

LEARN MORE ABOUT THE
DATA AVAILABLE

STEP
2

CHECK FOR YOUR
INSTITUTION'S AGREEMENT

STEP
3

REGISTER AS A RESEARCHER

STEP
4

VERIFY YOUR IDENTITY

STEP
5

COMPLETE *ALL OF US*
RESPONSIBLE CONDUCT OF
RESEARCH TRAINING

STEP
6

SIGN DATA USER CODE OF
CONDUCT



allof-us.org/Register

Stay in touch to learn more



Sign up for our bimonthly newsletter

Subscribe to Research Roundup

Stay up-to-date on the latest news and insights from the *All of Us* Research Hub through our bimonthly email newsletter.



allof-us.org/RRSignup

Thank You!



ResearchAllofUs.org



National Institutes
of Health

AllofUs.nih.gov



@AllofUsResearch
#JoinAllofUs