The *All of Us* Research Program

**Sheri Schully, Deputy Chief Medical and Scientific Officer, *All of Us*, NIH**

PMI Summit and 21st Century Cure Act

My hope is that this becomes the foundation, the architecture whereby 10 years from now we can look back and say we that we have revolutionized medicine.

- President Obama, February 25, 2016

The 21st Century Cures Act

The Cures Act, or the 21st Century Cures Act 🔗 **pdf**, passed overwhelmingly in both the U.S. House of Representatives and Senate with strong bipartisan support, and was signed into law on December 13, 2016. The legislation provides NIH with critical tools and resources to advance biomedical research across the spectrum, from foundational basic research studies to advanced clinical trials of promising new therapies.

Signed into law December 2016

Funded four highly innovative scientific initiatives:

1. The *All of Us* **Research Program**
2. The *Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative*
3. The *Cancer Moonshot*ᔆᴹ
4. The *Regenerative Medicine Innovation Project*

All of Us Mission

Nurture partnerships for decades with at least a **million participants** who reflect the diversity of the U.S.

Our mission Accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care for all of us.

Deliver one of the largest, richest, biomedical datasets of its kind that is broadly available and secure

Catalyze an ecosystem of **communities, researchers, and funders** who make *All of Us* an **indispensable** part of health research

Current protocol

Enroll, Consent and Authorize EHR
- Recruiting 18+ years old initially; plan to include children in future
- Online, interactive consent
- Includes authorization to share EHR data

Answer Surveys
- The Basics
- Overall Health
- Lifestyle
- Health Care Access & Utilization
- Personal and Family Medical History
- Social Determinants of Health
- NEW:
- Emotional Health History and Well-Being
- Behavioral Health and Personality

Provide Physical Measurements
- Blood pressure
- Heart rate
- Height
- Weight
- BMI
- Hip circumference
- Waist circumference

Provide Biosamples
- Blood (or saliva)
- Urine specimen
- Biosamples will be stored at the program's biobank

Share data from Wearables/Digital Apps
- Share data from wearable fitness devices, starting with Fitbit
- *Fitbit data may include physical activity, step counts, heart rate, and sleep data*

Genomics Plans

Infrastructure

Three Genome Centers: to generate genotyping & whole genome sequencing for one million or more participants

Genetic Counseling Resource: to support the responsible return of information to interested participants
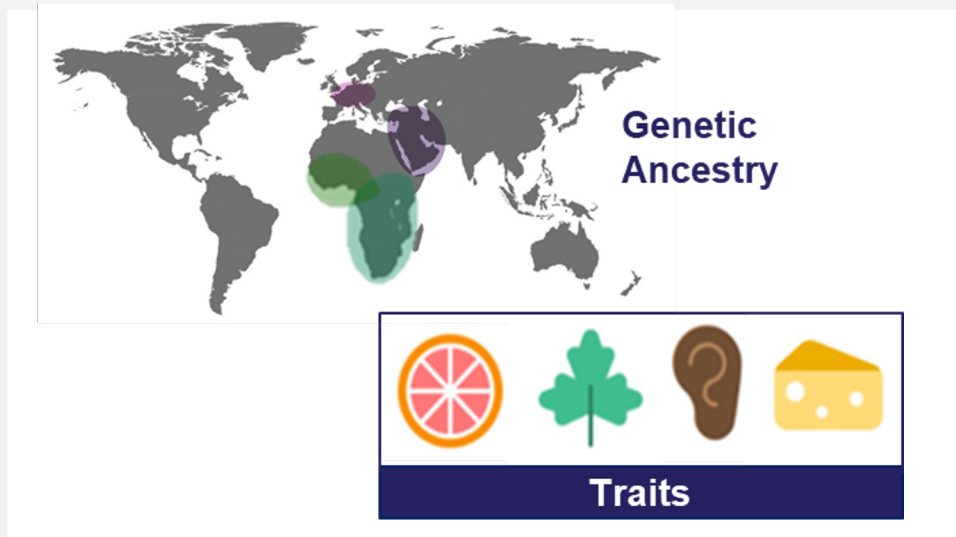
Return of Information

Over time, the program anticipates providing several kinds of information of interest to participants:

- Traits and ancestry
- Drug-gene interactions
- Genetic findings connected with risk of certain diseases

Genetic Return of Results: Engagement and Health-Related

Engagement Genetics

Genetic Ancestry

Traits

- Genetics Engagement Module (GEM)
- Genotyping
- Shared UI with Color Health

color

Health-Related Genetics

Hereditary Disease Risk
- ACMG v2.0
- Genome sequencing

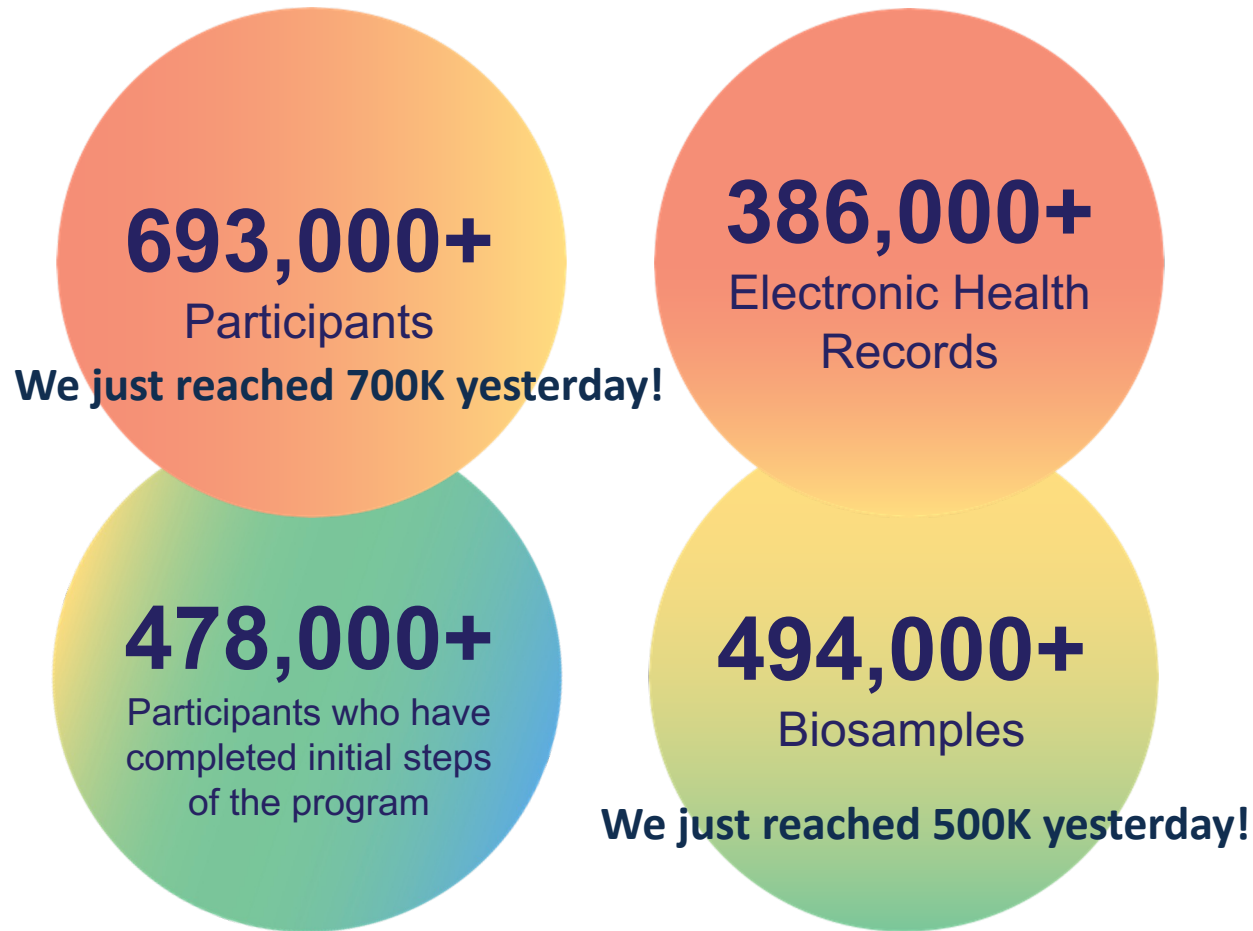**Medicine and your DNA**
- Pharmacogenetics
- Genome sequencing

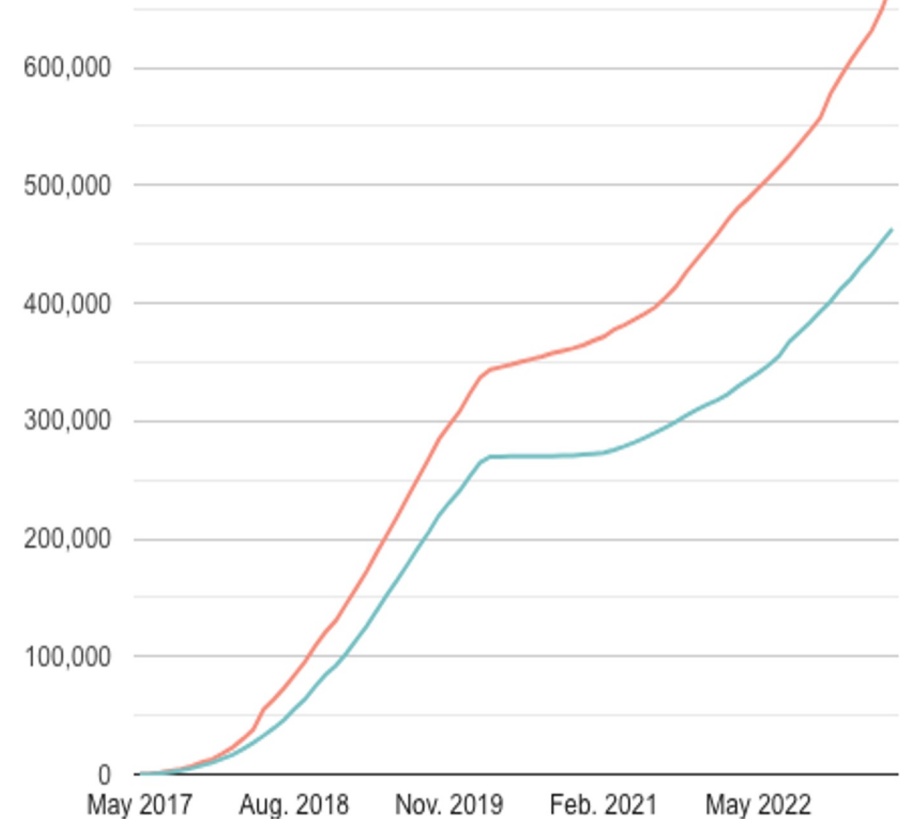**FDA** Investigational Device Exemption (IDE) issued July 9, 2020

All of Us Data

Enrollment data snapshot

693,000+
Participants

We just reached 700K yesterday!

386,000+
Electronic Health Records

478,000+
Participants who have completed initial steps of the program

494,000+
Biosamples
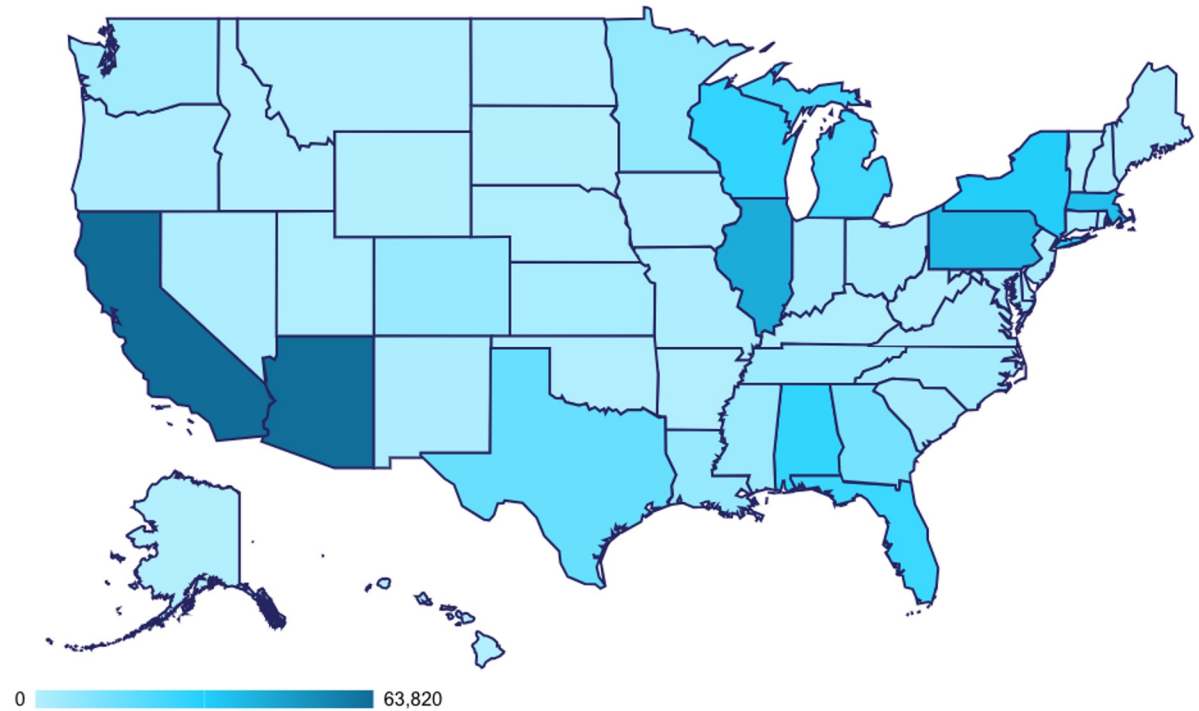
We just reached 500K yesterday!

Enrollment Numbers

Legend:
— Participants
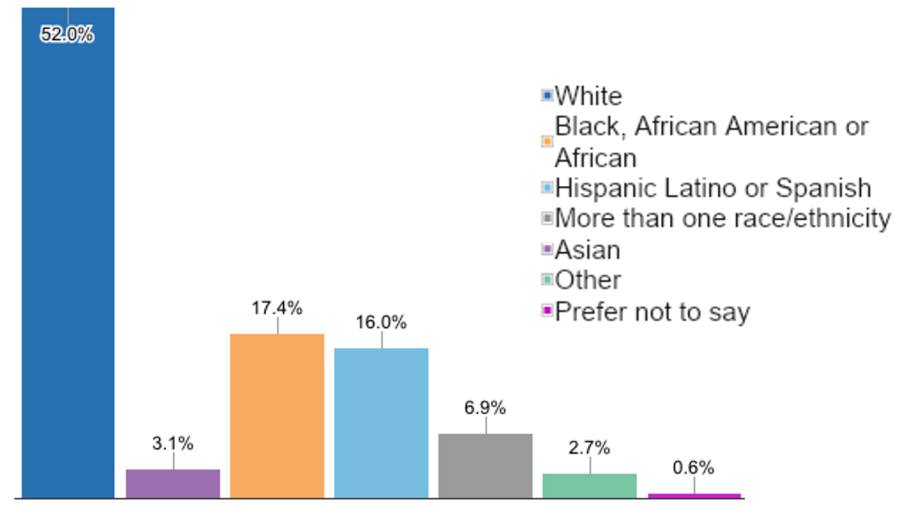— Participants who have completed the initial steps of the program

A bit about *All of Us* participants

80% of *All of Us* participants are underrepresented in biomedical research

Race and Ethnicity

- White — 52.0%
- Black, African American or African — 17.4%
- Hispanic Latino or Spanish — 16.0%
- More than one race/ethnicity — 6.9%
- Asian — 3.1%
- Other — 2.7%
- Prefer not to say — 0.6%

Age

Age	%
18-29	9.2%
30-39	14.7%
40-49	14.5%
50-59	18.0%
60-69	21.7%
70-79	16.2%
80-89	4.9%
89+	0.6%

ResearchAllofUs.org

Data as of August 2023
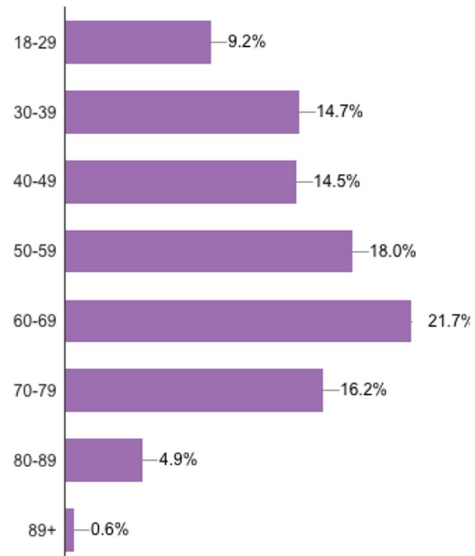
Data Types and Sources

Version 7- Released April 2023

413,350+
Survey Responses

337,500+
Physical Measurements

312,900+
Genotyping Arrays

287,000+
Electronic Health Records

245,350+
Whole Genome Sequences

15,600+
Fitbit Records
NEW! Sleep Data
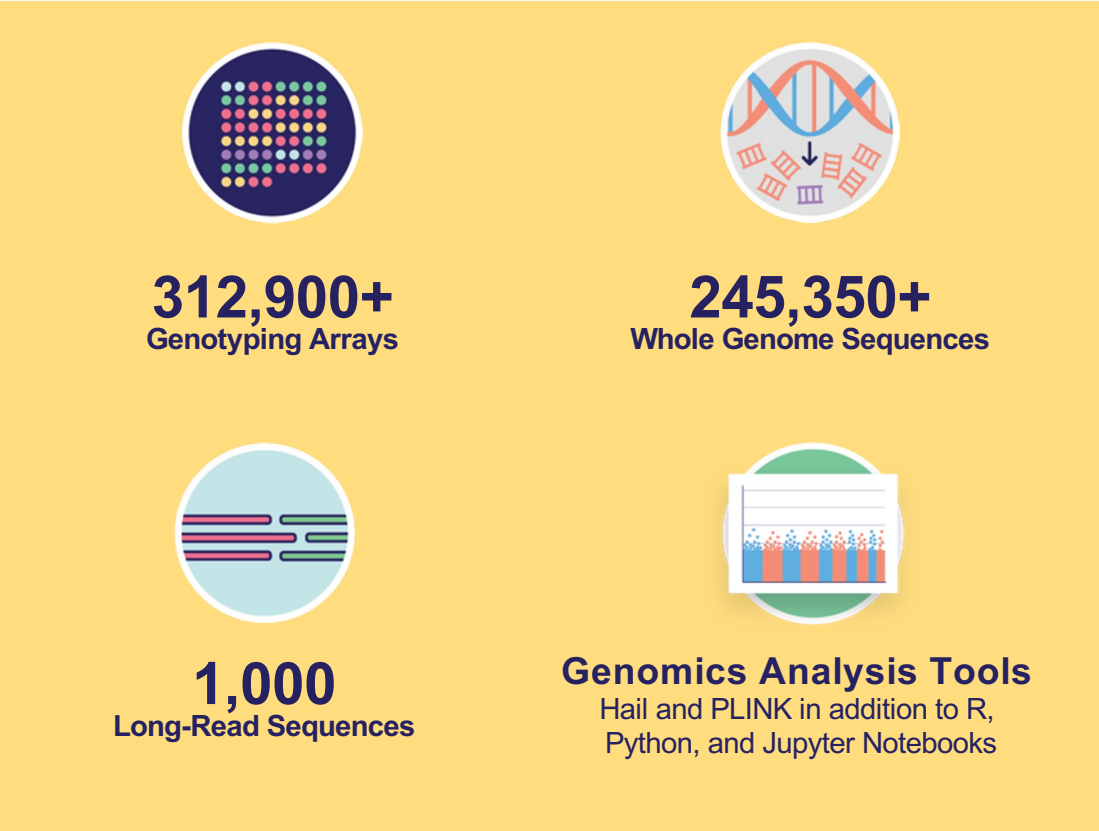
11,350+
Structural Variants
NEW! In 2023

1,000+
Long-Read Sequences
NEW! In 2023

Genomics data available

All of Us is the **largest set of whole genome sequences widely available for research.***

The whole genome sequence dataset includes variation at more than **1 billion** locations, which is nearly **one-third** of the entire human genome.
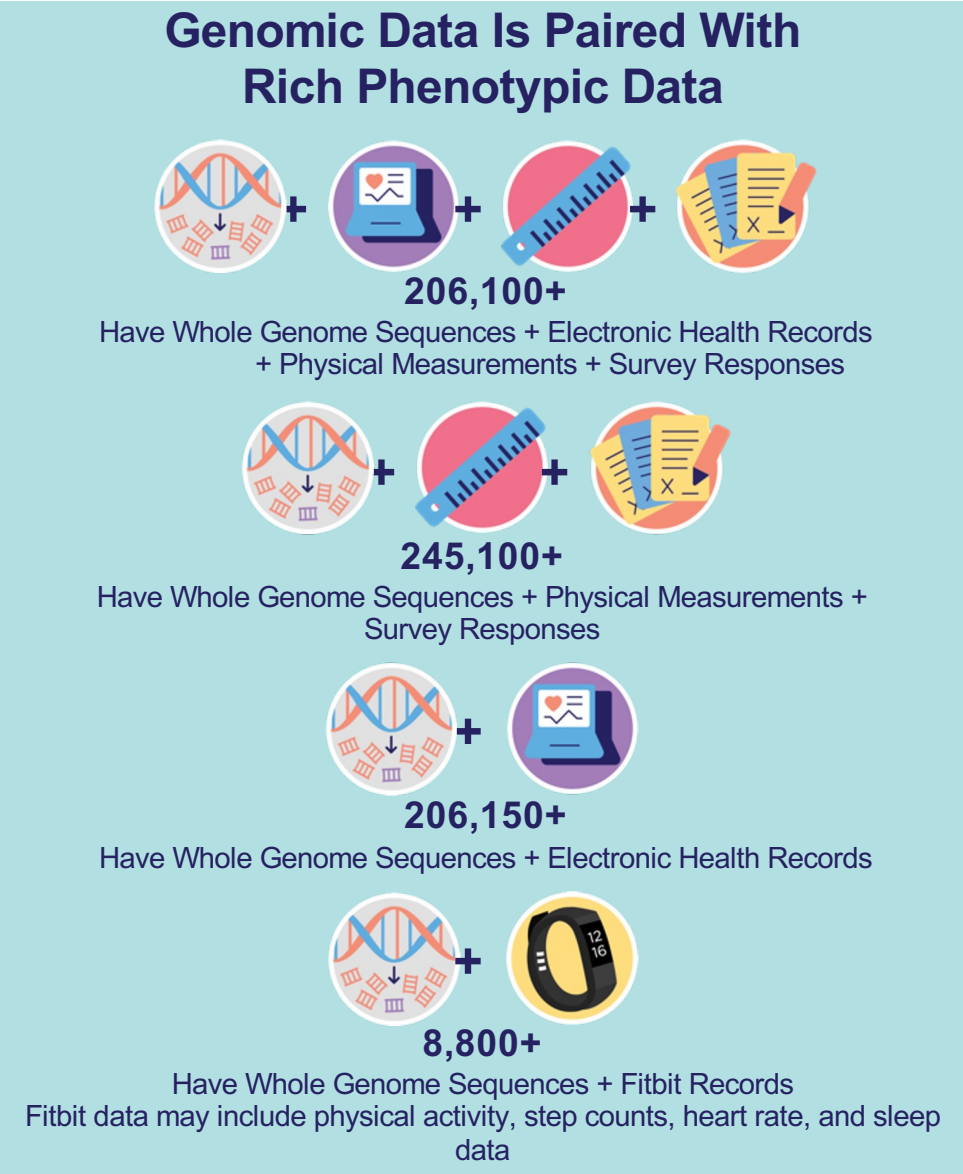
312,900+
Genotyping Arrays

245,350+
Whole Genome Sequences

1,000
Long-Read Sequences

Genomics Analysis Tools
Hail and PLINK in addition to R, Python, and Jupyter Notebooks

Only available in the Controlled Tier

Genomic Data Is Paired With Rich Phenotypic Data

206,100+
Have Whole Genome Sequences + Electronic Health Records + Physical Measurements + Survey Responses

245,100+
Have Whole Genome Sequences + Physical Measurements + Survey Responses

206,150+
Have Whole Genome Sequences + Electronic Health Records

8,800+
Have Whole Genome Sequences + Fitbit Records
Fitbit data may include physical activity, step counts, heart rate, and sleep data

All of Us Research: How It Works
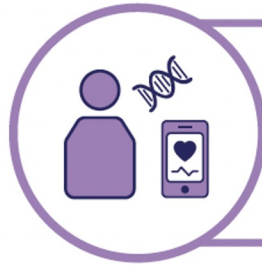
1. **Participants share their data** with the *All of Us* Research Program through multiple sources. These data are sent to a secure cloud environment, managed by the Data and Research Center.

2. **Participant data** is received and funneled through a **curation pipeline*** within a secure repository that connects to the Research Hub tools.
   * researchallofus.org/data-tools/data-sources

3. **Anyone can visit the Research Hub** to learn more about the types of data *All of Us* makes available for research. The **Survey Explorer** [1] and **Data Browser** [2] offer more information about the unique data elements and let visitors browse aggregated participant data.
   [1] researchallofus.org/data-tools/survey-explorer
   [2] databrowser.researchallofus.org

4. **Researchers register*** for access to the Researcher Workbench to analyze data.
   * researchallofus.org/register

5. **Registered researchers** in the Researcher Workbench can create research projects using collaborative workspaces, cohort-building tools, interactive notebooks, and more.

6. **Research underway** can be viewed on the **Projects Directory**.[1] Publications related to *All of Us* Research Program data are posted on the **Publications page**.[2]
   [1] researchallofus.org/research-projects-directory
   [2] researchallofus.org/publications

Research Hub Website

ResearchAllofUs.org

Goal: To provide more information around program data and tools.

- **Data Snapshots**: broad cohort metrics
- **Data Browser**: interactive tool
- **Survey Explorer**: source information for participant-provided information
- **Researcher Workbench**: restricted cloud-based platform designed to execute rapid, hypothesis-driven research
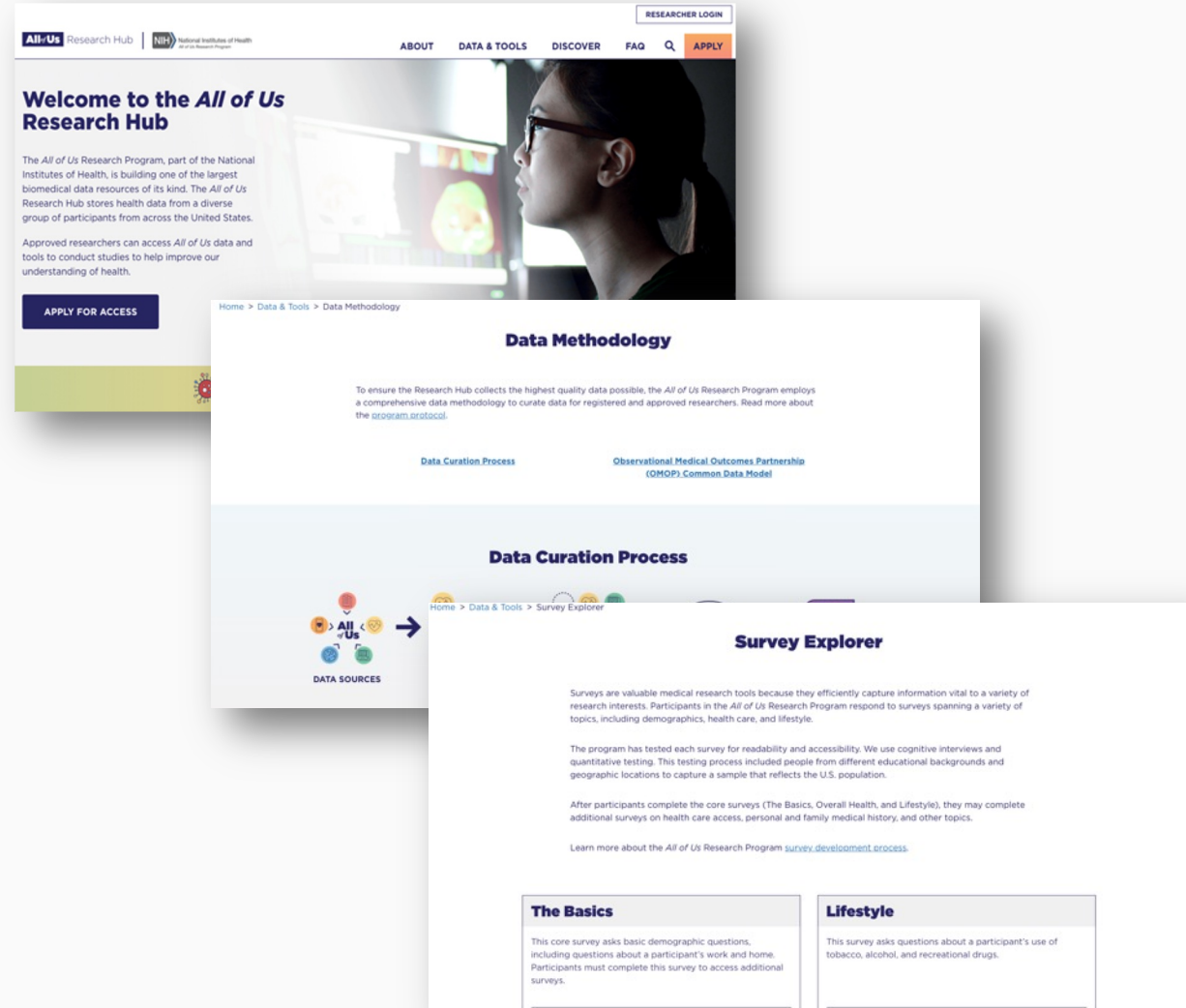
Powerful analytical tools & user support

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
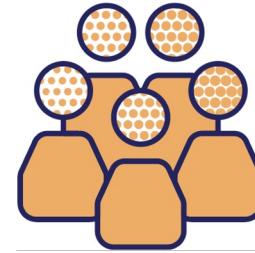
**COHORT
BUILDER**

USES:
 Cohort creation

**USER SUPPORT
HUB**

USES:
 Workbench training and
 resources
 Articles on Data and Tools
 Live office hours

Broadly accessible to researchers across stages and settings

Our Researchers

6,400+
Registered Researchers
across a range of institutional roles and career stages

550+
Institutions

Researchers **nationwide** are using the **Researcher Workbench**

Bar chart:
- Academic Institutions*: 350+
- Not-For-Profit Organizations: 100+
- Health Care Institutions: 70+
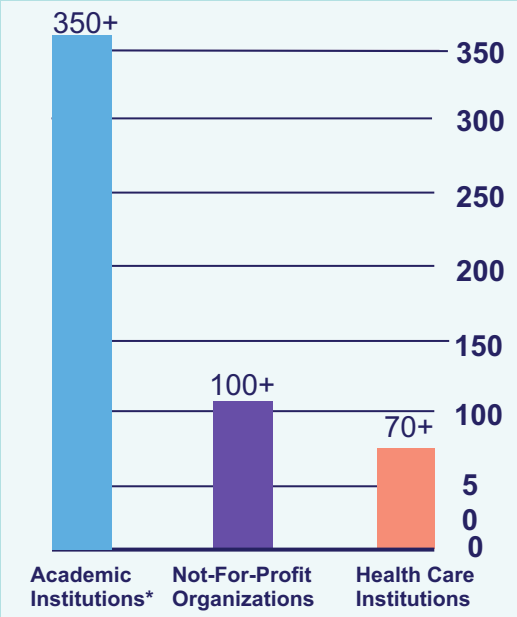
*Includes 90+ Historically Black Colleges and Universities and Hispanic-Serving Institutions.

Research Currently Underway

6,300+
Active projects

175+
Publications in peer-reviewed journals

Top conditions being studied
- Cardiovascular disease
- Hypertension
- Mental Health
- Cancer
- Diabetes

Figures accurate as of April 2023

15

Advantages and limitations: Data

Advantages

- DIVERSITY of participants and data (75% of participants are from communities underrepresented in biomedical research, including 45% of whom identify with racial and ethnic minority groups)

- DIMENSIONALITY: Unparalleled ability to make associations through our integrated dataset by cross-referencing participants' genes, environment, and life circumstances

- GRANULARITY OF PARTICIPANT-PROVIDED DATA enables better understanding of race, ethnicity, gender, and sexual orientation (based on how participants self-identify)

- BROADLY ACCESSIBLE thanks to cloud-based platform; you don't need expensive equipment to access the data

- NEW DATA RELEASED REGULARLY: Ongoing data releases as enrollment continues

Data as of April 2023

Limitations

- **NOT NATIONALLY REPRESENTATIVE:** All of Us oversamples for those who have traditionally been underrepresented in biomedical research

- **LIMITED ENVIRONMENTAL DATA:** Additional data from surveys and connections to external datasets forthcoming

- **POTENTIAL LEARNING CURVE FOR NEW INTERACTIVE ANALYSIS TOOLS** (Hail & PLINK) supporting R & Python and Command Line

Advantages and limitations: Research tools

Advantages

- POWERFUL ANALYTIC CAPABILITIES: Advanced computing potential enables the simultaneous exploration of multiple questions on a powerful scale supported by R and Python, the most widely used open-source platforms for statistical analysis, as well as HAIL and Plink, both commonly used for genomic analysis.

- TEAM SCIENCE APPROACH: The Researcher Workbench provides the data, tools, and opportunity to come together with other researchers from different disciplines around project-specific data analysis (through collaborative workspaces, cohort-building tools, interactive notebooks, and more).

- AVAILABILITY OF FEATURED WORKSPACES AND SAMPLE COHORTS: For replicability and to help researchers get started.

Limitations

- **LEARNING CURVE WITH DATA ANALYSIS TOOLS:** Data analysis requires proficiency in Jupyter notebooks and ability to code in R, Python, HAIL and Plink. Some institutions and teams may come to the platform without a background in these tools.

- **CAPACITY:** Smaller research teams may find that capacity constraints (e.g., cost, proficiency in coding tools, dependency on data scientists and/or partnerships) impact their ability to make full use of the All of Us platform without collaborators.

- **ADDITIONAL COSTS ASSOCIATED WITH ANALYSIS OF GENOMIC DATA:** While there is no cost to access the Researcher Workbench, there are computational costs for analyses incurred through Google Cloud Platform. The program provides $300 in initial credits to registered researchers to get them started.

17

Data as of April 2023

All of Us Community and Provider Partner Network *(as of April 2023)*

All of Us Consortium Members *(beyond community partners, as of April 2023)*

The Participant Center

Scripps Research Translational Institute · BlueCross BlueShield · SENSIS · SageBionetworks · waves · DENVER HEALTH FOR LIFE'S JOURNEY · MAYO CLINIC · Bloodworks Northwest · BLOOD ASSURANCE · OurCEL SOLUTIONS

careevolution · WebMD · Montage Marketing · South Texas Blood & Tissue · SunCoast Blood Centers · Sync for Science · Quest Diagnostics · San Diego Blood Bank · fitbit · Mary's Center

Communications & Engagement

WONDROS · PYXIS PARTNERS · RTI INTERNATIONAL · THE UNIVERSITY OF UTAH

HPO Network
(Health Care Provider Organizations)

HPO Lite

UTTyler · VCU Health · Washington University in St.Louis

RMCs

All of Us California

UC San Diego Health · UCI Health · UC DAVIS HEALTH · UCSF · ECRMC · Cedars Sinai · Keck School of Medicine of USC

Illinois Precision Medicine Consortium

Northwestern Medicine · THE UNIVERSITY OF CHICAGO · NorthShore University HealthSystem · RUSH · THE UNIVERSITY OF ILLINOIS COLLEGE OF MEDICINE · UI Health · COOK COUNTY HEALTH

All of Us New England

Mass General Brigham · BOSTON MEDICAL CENTER

Trans America Consortium

Henry Ford HEALTH SYSTEM · Spectrum Health · Essentia Health · BaylorScott&White HEALTH · RELIANT MEDICAL GROUP Part of OptumCare® · HealthPartners Institute

New York City Consortium

COLUMBIA · Weill Cornell Medicine · NYC HEALTH + HOSPITALS | Harlem · NewYork-Presbyterian

All of Us Southern Network

UAB MEDICINE · USA HEALTH · UAB HEERSINK The University of Alabama at Birmingham · University Medical Center · THE UNIVERSITY OF ALABAMA · Cooper Green Mercy HEALTH SERVICES AUTHORITY · TUSKEGEE UNIVERSITY · The University of Mississippi MEDICAL CENTER · LSU Health NEW ORLEANS · Tulane University

All of Us Southeast Enrollment Center

UHealth UNIVERSITY OF MIAMI MILLER SCHOOL OF MEDICINE · EMORY UNIVERSITY · MOREHOUSE SCHOOL OF MEDICINE · UF UNIVERSITY of FLORIDA · THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

All of Us Puerto Rico

cossma · CENTRO COMPRENSIVO DE CANCER

All of Us Wisconsin

Marshfield Clinic Health System · WISCONSIN UNIVERSITY OF WISCONSIN-MADISON · Froedtert & MEDICAL COLLEGE OF WISCONSIN · GUNDERSEN HEALTH SYSTEM

All of Us Pennsylvania

University of Pittsburgh

University of Arizona and Banner Health

THE UNIVERSITY OF ARIZONA · Banner Health · MARIPOSA COMMUNITY HEALTH CENTER

FQHCs (Federally Qualified Health Centers)

Cherokee Health Systems · SAN YSIDRO HEALTH · WINNEBAGO COAST COMPREHENSIVE HEALTH CENTER · Community Health Center, Inc. · Cooperative Health · Sun River Health

VA Medical Centers

VA · U.S. Department of Veterans Affairs Veterans Health Administration

Nutrition for Precision Health (NPH)

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL · UNC NUTRITION INSTITUTE · AdventHealth · Mass General Brigham

---

Participant Technology Systems Center (PTSC)

vibrent

Biobank

MAYO CLINIC

Data & Research Center (DRC)

VANDERBILT UNIVERSITY MEDICAL CENTER · BROAD INSTITUTE · verily · COLUMBIA · SCHOOL OF PUBLIC HEALTH UNIVERSITY OF MICHIGAN · Northwestern Medicine · UTHealth The University of Texas Health Science Center at Houston

Genomics Partners

Baylor College of Medicine DEPARTMENT OF MOLECULAR & HUMAN GENETICS · JOHNS HOPKINS MEDICINE · UTHealth The University of Texas Health Science Center at Houston · HudsonAlpha INSTITUTE FOR BIOTECHNOLOGY · BROAD INSTITUTE · color · Mass General Brigham · UNIVERSITY of WASHINGTON

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

Create an *All of Us* account

Get started today

STEP 1
CONFIRM YOUR INSTITUTION'S AGREEMENT

STEP 2
CREATE AN ACCOUNT AND VERIFY YOUR IDENTITY

STEP 3
COMPLETE THE MANDATORY TRAINING

STEP 4
SIGN THE DATA USER CODE OF CONDUCT (DUCC)

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.

allof-us.org/RRSignup