Presenters

Ashley Greene, MLIS
Senior Research Projects Coordinator
ashley.e.green@vumc.org

Objectives

At the end of this session, you should be able to:

- Enhance knowledge of the *All of Us* dataset and Researcher Workbench
  - Describe the *All of Us* Research Program
  - Describe the data offered in the public data browser and the Researcher Workbench
  - Describe how to access the data
  - Identify tools on the public and workbench sites that are helpful for data exploration and analysis
- Understand the *All of Us* Researcher Workbench's value to your campus community
  - Understand how to leverage your role as a librarian in campus-wide use of the *All of Us* Research Hub
  - Identify research fields that will benefit from access to the data

Making Health Discoveries Possible

The *All of Us* Program wouldn't be possible without the generosity of our participants and the dedication of our researchers to enable health discoveries.

Inviting

1 Million

or more people
across the
United States

Data in the Researcher Workbench
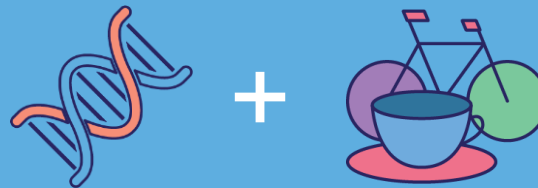
329,000+
Participants

80%
are from underrepresented
communities

50%
are racial and
ethnic minorities

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

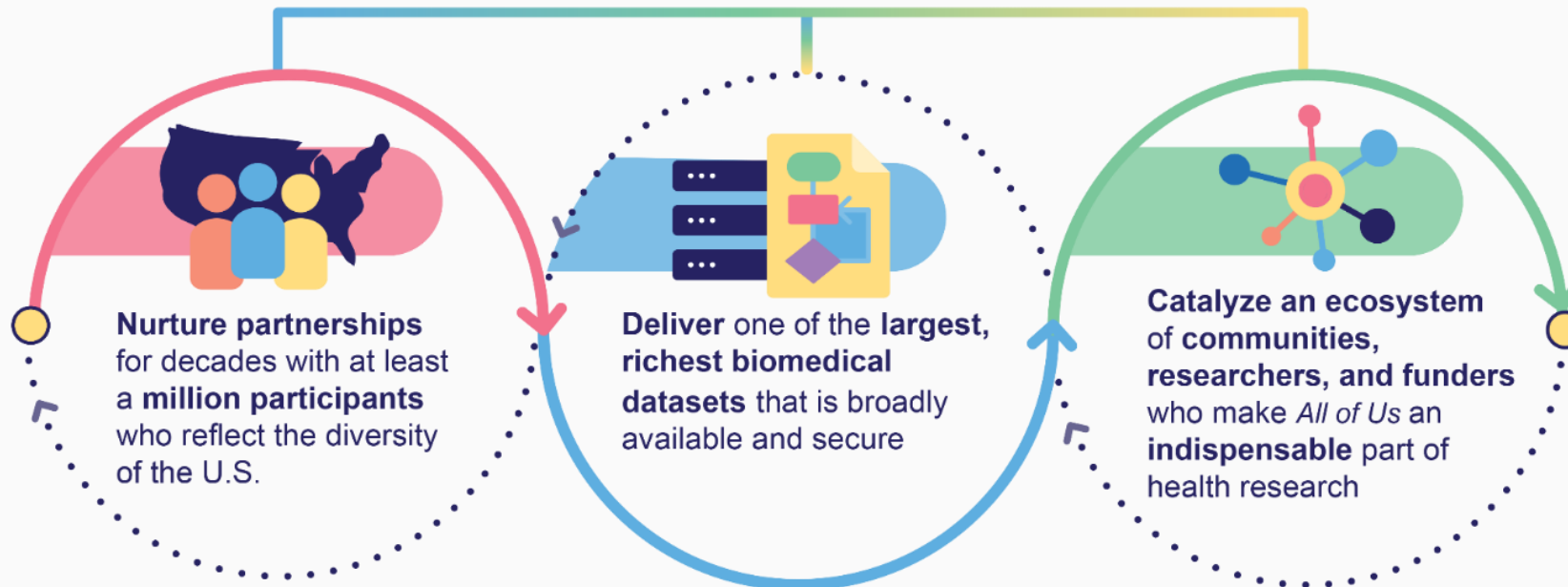Easily **accessible** to any researcher with a secure internet connection

Follows participants as they move, age, and grow

Our Mission

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

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

Deliver one of the **largest, richest biomedical datasets** 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

Made possible by a team that maintains a culture built around the program's core values

Participants included in All of Us research data are diverse.

Underrepresented in Biomedical Research (UBR) Categories	Curated Data (% out of 413,457 participants)
At least one UBR	74%
Non-white race or Hispanic/Latino ethnicity	43%
Age >= 65	24%
Less than GED	9%
Annual Income <=$25k	25%
Sexual and Gender Minorities	10%
Disability	10%

Summary of Data Available in the Researcher Workbench

413,350+
Survey Responses
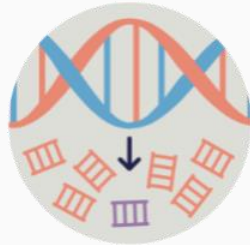
337,500+
Physical Measurements

312,900+
Genotyping Arrays

287,000+
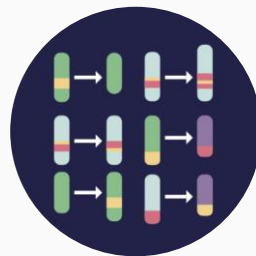Electronic Health Records

245,350+
Whole Genome Sequences

15,600+
Fitbit Records
NEW! Sleep Data

11,350+
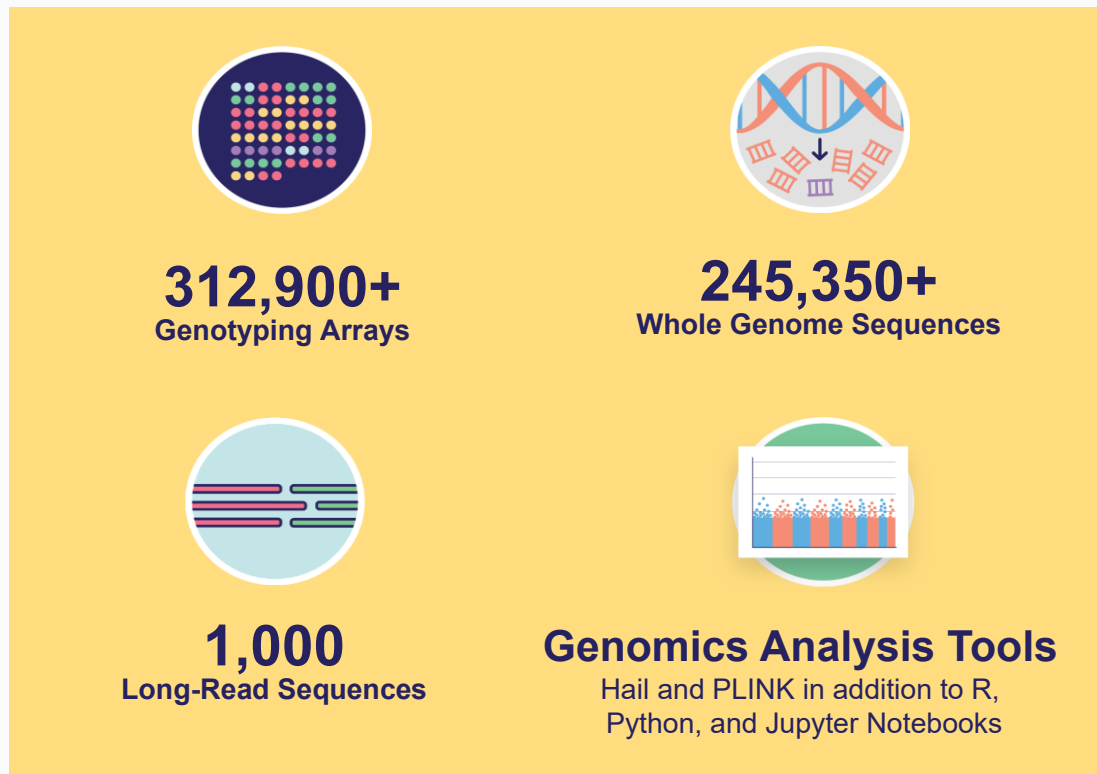Structural Variants
NEW! In 2023

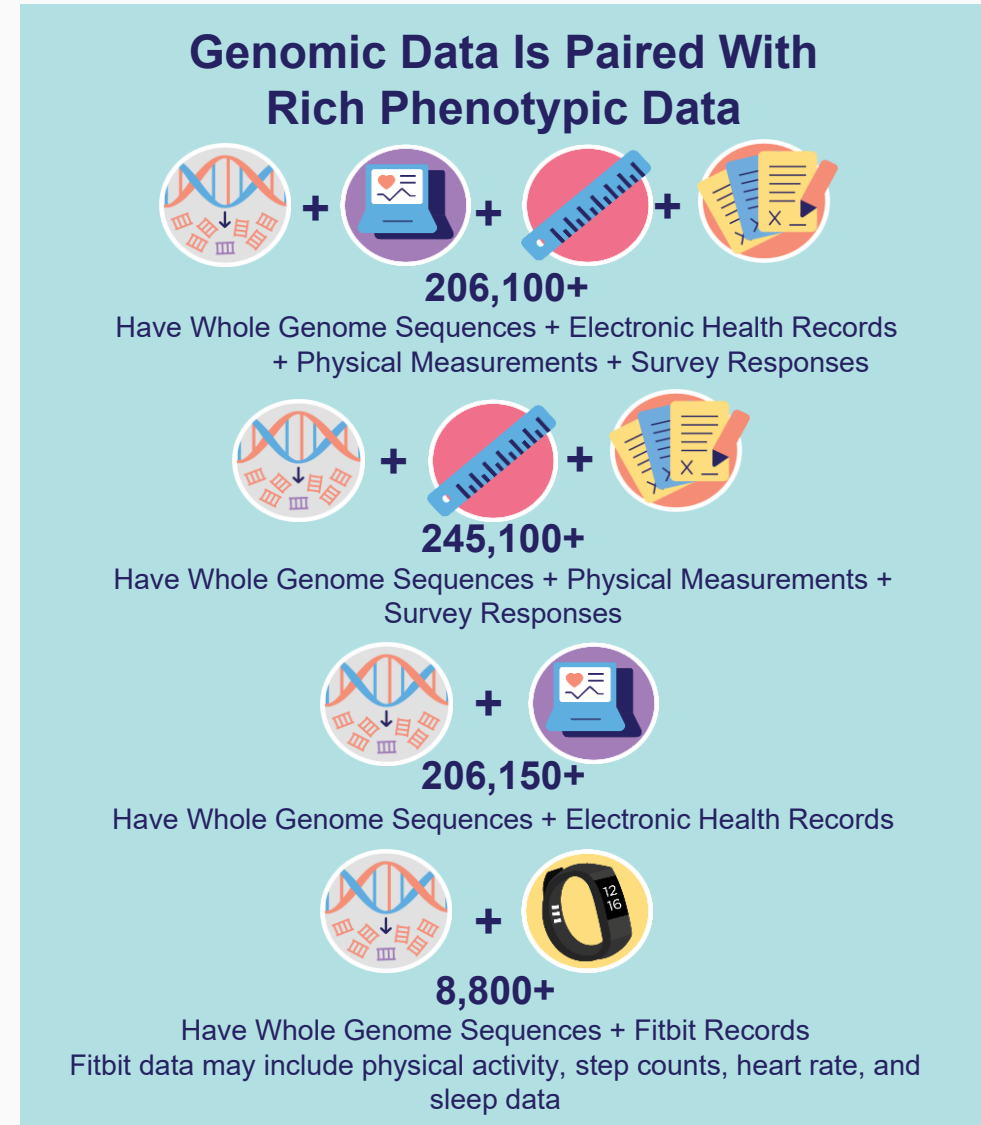**1,000+**
Long-Read Sequences
NEW! In 2023

Data as of April 2023

Including Genomic Data

As of April 2023, the *All of Us* Researcher Workbench contains 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
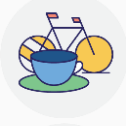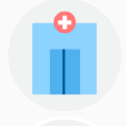
Data as of April 2023

Data: Surveys

1. The Basics
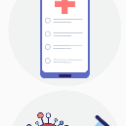2. Overall Health
3. Lifestyle
4. Health Care Access & Utilization
5. Personal Health History and Family Medical History (Revised)
6. Social Determinants of Health (SDOH)
7. COVID-19 Participant Experience (COPE)
8. COVID-19 Vaccines Minute Surveys

Visit the Survey Explorer on ResearchAllofUs.org for more information on survey data

Data: Social Determinants of Health Survey

By 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.

Social Determinants of Health (SDOH) Survey data

Neighborhood Safety · Neighborhood Environment · Social Support · Discrimination · Loneliness · Health Care Discrimination · Housing · Food Security · Stress · Spiritual Experiences · Religious Attendance · English Proficiency

from 117,750+ responses

Data as of April 2023

Data: Cope Survey

All of Us released its recurring COVID-19 Participant Experience (COPE) survey to learn more about changes in participants' daily life, health, and well-being throughout the pandemic. Registered researchers can access the data in the Researcher Workbench and use it as part of an integrated dataset to explore potential associations across a variety of factors, from genetics to geography to mental health.

COPE Survey Data Currently Available in the Researcher Workbench

Nearly **100,200** participants responded to one or more of the **six** COPE surveys administered between May 2020 and February 2021

Topics Covered

- Social Distancing Experiences
- COVID-19 Related Symptoms
- COVID-19 Related Testing
- COVID-19 Related Treatment
- COVID-19 Related Impact
- General Well-Being
- Social Support
- Anxiety
- Mood
- Stress
- Physical Activity
- Loneliness
- Substance Use
- Resilience
- Discrimination
- Vaccine Perceptions

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
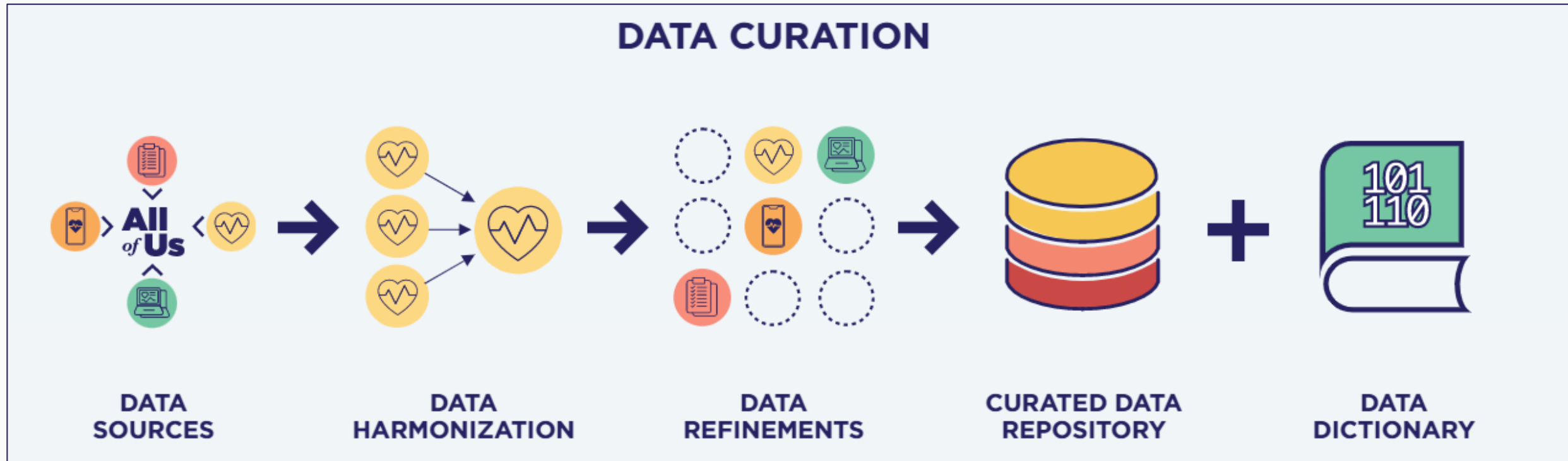- **94%** have someone to love and make them feel wanted
- **95%** have someone to have a good time with

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

*Participants could select multiple symptoms.

All of Us Research Data Pipeline

DATA CURATION

DATA SOURCES → DATA HARMONIZATION → DATA REFINEMENTS → CURATED DATA REPOSITORY + DATA DICTIONARY

Data: Access Flowchart and Tools

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 participant identifiers removed. Public resources include:

- **Data Snapshots:** Aggregated, public-facing overviews of participant characteristics and data types
- **Data Browser:** Interactive preview into the *All of Us* dataset through public-facing aggregate data
  - Currently includes participant-provided survey responses, physical measurements, data from EHRs and wearables, and genomic data
- **Survey Explorer:** Details the questions included in each of the surveys
- **Research Projects Directory:** Descriptions of each research project within the Researcher Workbench

Researcher Workbench

Registered Tier

Registered researchers can access in-depth data and a variety of research tools to conduct a range of studies.

Surveys

Electronic Health Records

Physical Measurements

Wearables

Data have been processed to protect participant privacy.

Controlled Tier

Registered researchers with amended institutional agreements can access all of the data in the Registered Tier plus additional and expanded data types, including genomic data, real dates of health events, ICD codes, more granular demographic data, and more.

Genomics

Health and Lifestyle surveys

Counts in the Public, Registered, and Controlled Tiers may vary based on the lag time associated with the data curation process under the program's rigorous privacy protections.

Data Release Timeline

2019
Data Browser
(launched May 2019)

Interactive tool available to the public that provides summary statistics from the program's growing database, allowing researchers to understand the characteristics of our participant population, **explore the data types available and plan research questions**

2020
Researcher Workbench
(launched May 2020)

Released of the first version of the *All of Us* dataset and the first set of researcher tools for beta testing

- **Survey** data
- **Physical measurement** data
- **EHR** data

2021
Expansion of data and tools
(throughout 2021)

Expanded of the *All of Us* dataset with additional survey, measurement, and EHR data.

- **Start of genotyping** and whole genome sequencing
- **Initial return of results** to participants

>1,000K researchers registered

2022
Controlled Tier & Genomics
(ongoing)

Launch our Controlled tier, available to approved researchers who have taken additional steps and training to access:

- **Genomics** data
- **Individual and granular-level** information

Available in Registered and Controlled tiers, with basic information available via our public data browser:

- **ACS data linkage**
- **New surveys**

2023 and *beyond*
Additional data types and tools
(ongoing)

Expect to enroll 1M participants within 5-7 years of launch and continue to expand the *All of Us* dataset with new data releases, including over time:

- **Assay data**
- **Additional wearables data**
  - Fitbit sleep data
  - Apple Healthkit
- **New surveys** (e.g., Mental Health & Wellbeing)
- **New Linkages to external data sources**
- **Expanded researcher access**
- **Ancillary study data**

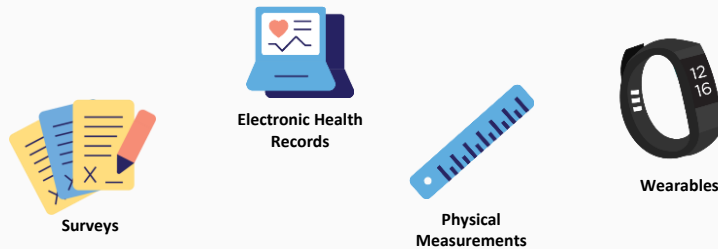**Release new versions of the Researcher Workbench**

Distinguishing between the Registered and Controlled Tier

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.

Surveys

Electronic Health Records

Physical Measurements

Wearables

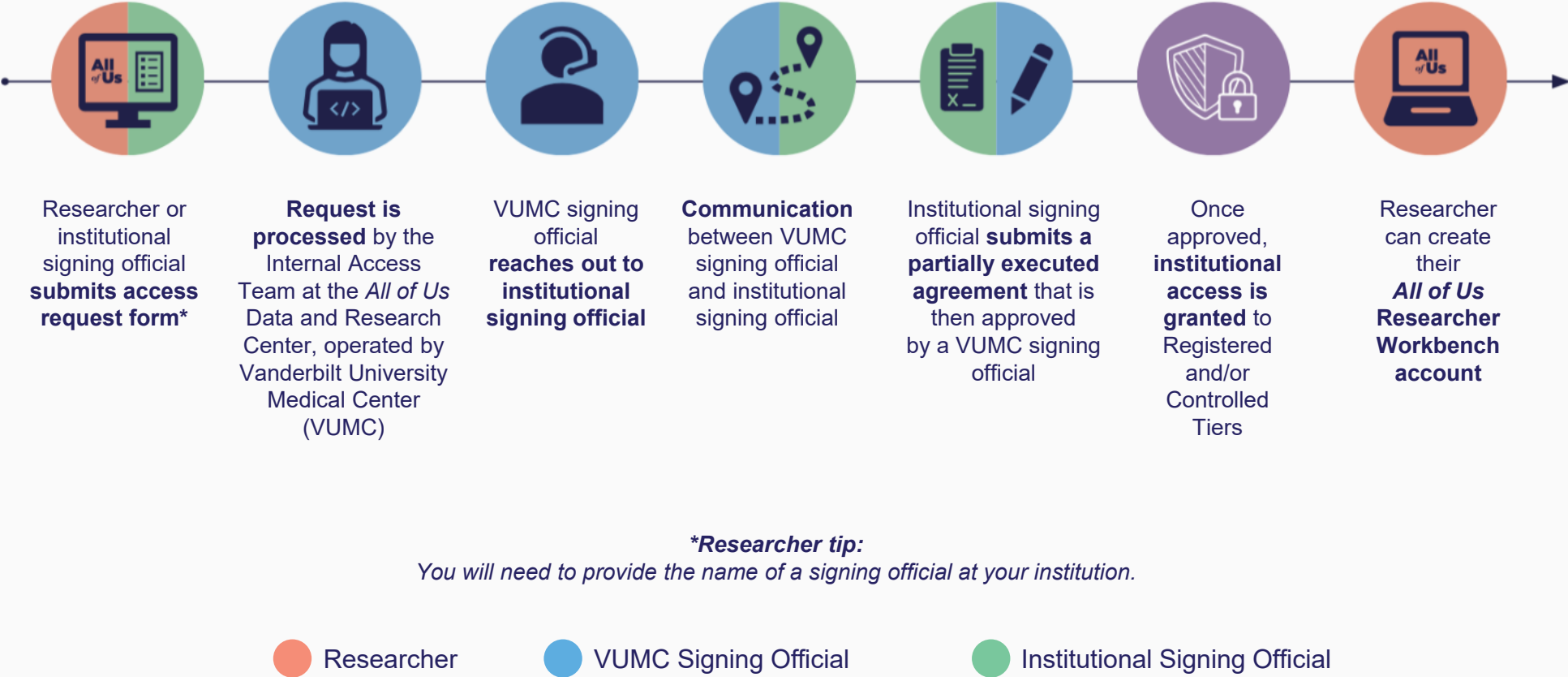*Data have been processed to protect participant privacy*

Controlled Tier

Registered users with amended institutional agreements can access **all of the data in the registered tier plus expanded and NEW data** including:

- Expanded demographics
- Responses to entire COPE Survey
- COVID-19 EHR data
- Unshifted event dates
- Genomic data derived from WGS (~100k participants) and array genotyping (>165k participants)

Genomics

How to Gain Institutional Access to the *All of Us* Dataset

The Data Use and Registration Agreement (DURA) Process

Researcher or institutional signing official **submits access request form***

Request is processed by the Internal Access Team at the *All of Us* Data and Research Center, operated by Vanderbilt University Medical Center (VUMC)

VUMC signing official **reaches out to institutional signing official**

Communication between VUMC signing official and institutional signing official

Institutional signing official **submits a partially executed agreement** that is then approved by a VUMC signing official

Once approved, **institutional access is granted** to Registered and/or Controlled Tiers

Researcher can create their *All of Us* **Researcher Workbench account**

*Researcher tip:
You will need to provide the name of a signing official at your institution.

⬤ Researcher   ⬤ VUMC Signing Official   ⬤ Institutional Signing Official

Submit a DURA Request

Visit the *All of Us* Research Hub to get started

https://allof-us.org/DURAForm

The Research Hub

Follow along at

https://www.researchallofus.org

Summary: Getting Your Library Involved

There are many ways to support research activities at your institution!

- Help a class with registered students use the cohort and dataset builder
- Collaborate on a research project
- Include *All of Us* in your Python or R instruction
- Provide research assistance to students and faculty on big data
- Promote the *All of Us* database for student projects and dissertations
- Find articles using the *All of Us* database
- Review the Research Projects Directory for research ideas
- Introduce the *All of Us* database to researchers from various disciplines (e.g., medicine, public health, nursing, social work).
- Share funding opportunities throughout campus
- Suggest a virtual demonstration for a campus group

Support Resources: Newsletter

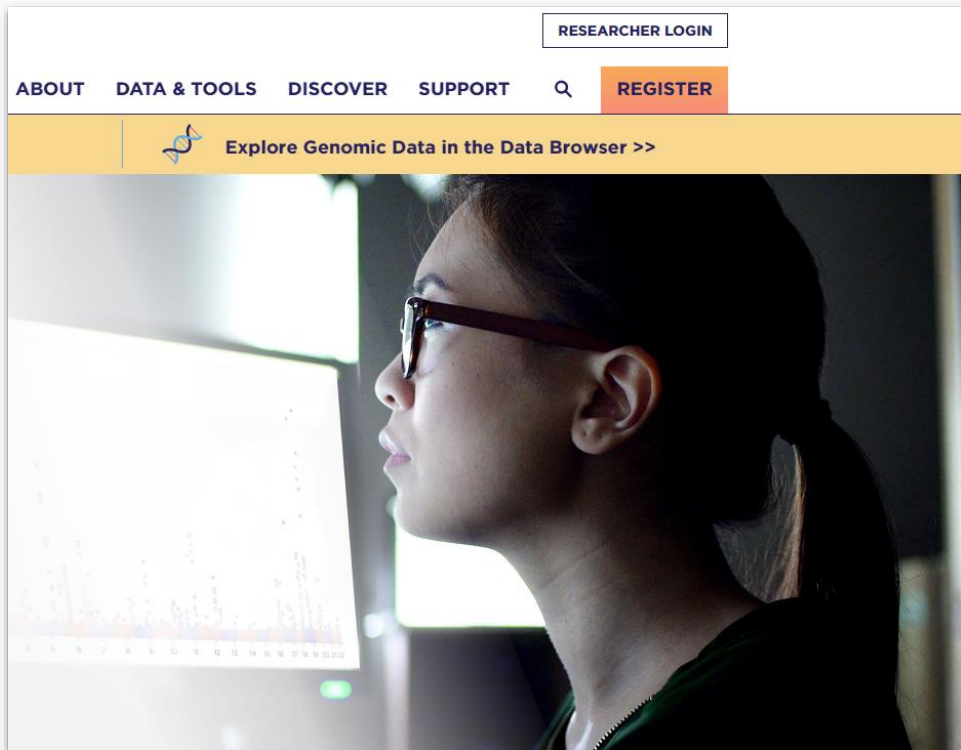## Sign up for our bimonthly newsletter

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.
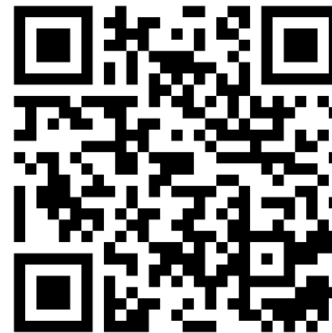

allof-us.org/RRSignup

Join our community of researchers!

Help us connect with other researchers!

Email **ashley.e.green@vumc.org** if you know of others interested in demos or training sessions

Sign up for our newsletter! **allof-us.org/RRSignup**

@AllofUsResearch
#JoinAllofUs