Journal of the American Medical Informatics Association, 2025, **32(1)**, 38–50
https://doi.org/10.1093/jamia/ocae270
Research and Applications

Research and Applications

Research for all: building a diverse researcher community for the *All of Us* Research Program

Rubin Baskir, PhD[1,*], Minnkyong Lee, PhD[1], Sydney J. McMaster, MPH[1], Jessica Lee, BS[1], Faith Blackburne-Proctor, MPA[1], Romuladus Azuine, DrPH, MPH, RN[1], Nakia Mack, MPH, MS[1], Sheri D. Schully, PhD[1], Martin Mendoza, PhD[1], Janeth Sanchez, PhD, MPH[1], Yong Crosby, MBA[1], Erica Zumba, MBA, MPH[1], Michael Hahn, BA[1], Naomi Aspaas, BSN, RN[1], Ahmed Elmi, MPH, CHES[1], Shanté Alerté, MPH[2], Elizabeth Stewart, MBA[3], Danielle Wilfong, MS, EdM[3], Meag Doherty, BS[1], Margaret M. Farrell, MPH, RD[1], Grace B. Hébert, PhD, MA[1], Sula Hood, PhD, MPH[4], Cheryl M. Thomas, CPA[5], Debra D. Murray, PhD[6], Brendan Lee, MD, PhD[6], Louisa A. Stark, PhD[7], Megan A. Lewis, PhD, MA[4], Jen D. Uhrig, PhD[4], Laura R. Bartlett, MLS[8], Edgar Gil Rico, MBA[9], Adolph Falcón, MPP[9], Elizabeth Cohn, PhD, NP, RN[10], Mitchell R. Lunn, MD, MAS, FACP, FASN[11], Juno Obedin-Maliver, MD, MPH, MAS, FACOG[12], Linda Cottler, PhD, MPH, FACE[13], Milton Eder, PhD[14], Fornessa T. Randal, EdD, MCRP[15], Jason Karnes, PharmD, PhD, BCPS[16], KiTani Lemieux, PhD[17], Nelson Lemieux Jr[18], Nelson Lemieux III, MS[18], Lilanta Bradley, PhD, MFT, CHES[19], Ronnie Tepp, MPP[20], Meredith Wilson, MPP[20], Monica Rodriguez, MS[20], Chris Lunt[1], Karriem Watson, DHSc, MS, MPH[1]

[1]All of Us Research Program, National Institutes of Health, Bethesda, MD 20817, United States, [2]National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, United States, [3]Data and Research Center, Vanderbilt University Medical Center, Nashville, TN 37232, United States, [4]RTI International, Research Triangle Park, NC 27709, United States, [5]Delta Research and Educational Foundation, Washington, DC 20009, United States, [6]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, United States, [7]Genetic Science Learning Center, University of Utah, Salt Lake City, UT 84112, United States, [8]National Library of Medicine, Washington, DC 20894, United States, [9]National Alliance for Hispanic Health, Washington, DC 20036, United States, [10]Feinstein Institutes for Medical Research, Manhasset, NY 11030, United States, [11]Department of Nephrology, Stanford University School of Medicine, Stanford, CA 94305, United States, [12]Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA 94305, United States, [13]Department of Epidemiology, University of Florida, Gainesville, FL 32611, United States, [14]Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, MN 55455, United States, [15]Asian Health Coalition, Chicago, IL 60605, United States, [16]Department of Pharmacy Practice and Science, University of Arizona, Tuscon, AZ 85721, United States, [17]Division of Basic Pharmaceutical Sciences, Xavier University of Louisiana, LA 70125, United States, [18]Seven Star Academy, Inc., New Orleans, LA 70114, United States, [19]Department Community Medicine and Population Health, University of Alabama, Tuscaloosa, AL 35487, United States, [20]Pyxis Partners, Washington, DC 20005, United States

*Corresponding author: Rubin Baskir, PhD, All of Us Research Program, National Institutes of Health, Bethesda, MD 20817, United States (rubin.baskir@nih.gov)

Abstract

Objectives: The NIH *All of Us* Research Program (*All of Us*) is engaging a diverse community of more than 10 000 registered researchers using a robust engagement ecosystem model. We describe strategies used to build an ecosystem that attracts and supports a diverse and inclusive researcher community to use the *All of Us* dataset and provide metrics on *All of Us* researcher usage growth.

Materials and Methods: Researcher audiences and diversity categories were defined to guide a strategy. A researcher engagement strategy was codeveloped with program partners to support a researcher engagement ecosystem. An adapted ecological model guided the ecosystem to address multiple levels of influence to support *All of Us* data use. Statistics from the *All of Us* Researcher Workbench demographic survey describe trends in researchers' and institutional use of the Workbench and publication numbers.

Results: From 2022 to 2024, some 13 partner organizations and their subawardees conducted outreach, built capacity, or supported researchers and institutions in using the data. Trends indicate that Workbench registrations and use have increased over time, including among researchers underrepresented in the biomedical workforce. Data Use and Registration Agreements from minority-serving institutions also increased.

Received: April 10, 2024; **Revised:** September 5, 2024; **Editorial Decision:** September 21, 2024; **Accepted:** October 16, 2024
Published by Oxford University Press on behalf of the American Medical Informatics Association 2024.

Discussion: *All of Us* built a diverse, inclusive, and growing research community via intentional engagement with researchers and via partnerships to address systemic data access issues. Future programs will provide additional support to researchers and institutions to ameliorate *All of Us* data use challenges.

Conclusion: The approach described helps address structural inequities in the biomedical research field to advance health equity.

Key words: biomedical research; training programs; stakeholder engagement; research personnel; capacity building.

Objective

The National Institutes of Health's (NIH's) *All of Us* Research Program (*All of Us*) is a historic effort to collect and study data from at least one million people living in the United States.[1] The program aims to accelerate precision medicine research by creating one of the world's largest and most diverse databases for health research. Researchers access *All of Us* data through the *All of Us* Researcher Workbench, a cloud-based analytics platform.[2] As of July 2024, more than 11 340 researchers have registered to use the platform since it launched in May 2020. Achieving this level of researcher registrations was due in part to the deliberate outreach to and engagement of researchers. *All of Us* Research Program defines researcher engagement as intentional, bidirectional relationship, and capacity building with researchers and their communities through ongoing information sharing, collaboration, consultation, and support. Researcher engagement at *All of Us* occurs across all career stages, institution types, and demographic categories with an emphasis on researchers from communities that have historically been underrepresented in the biomedical workforce (UBW researchers). *All of Us* Research Program considers researchers to be UBW if they self-identify with one or more of the following groups: racial and ethnic minority groups, lesbian, gay, bisexual, transgender, queer, intersex, asexual, aromantic, and other sexual and/or gender minority (LGBTQIA+) communities, persons living with disabilities, or those who are from a disadvantaged background (Table 1). In developing this definition, *All of Us* staff consulted the Notice of NIH's Interest in Diversity,[3] conducted an internal literature review on underrepresentation in the biomedical workforce, and considered input from program partners (Table 1).[4,5] In addition to developing a definition of UBW researchers to support all researchers using the *All of Us* dataset, the Researcher Engagement and Outreach (REO) Branch within the Division of Engagement and Outreach (DEO) at *All of Us* developed the *All of Us* researcher engagement ecosystem and adapted a social–ecological model of researcher engagement.[6–9] The adapted model assists the branch and its partners in strategic planning and the evaluation of researcher engagement.

Table 1. *All of Us* definition of researchers underrepresented in the biomedical workforce (UBW).

Diversity category	Underrepresented in the biomedical workforce (UBW)
Race and ethnicity	Individuals who identify as American Indian/Alaska Native; Black; African American or of African descent; Hispanic, Latino, or of Spanish origin; Middle Eastern or North African; Native Hawaiian or Pacific Islander; 2 or more races; or "none of these fully describe me" Individuals who identify as one of the following Asian racial identities: Cambodian, Filipino, Hmong, Japanese, Lao, Pakistani, or Vietnamese
Gender identity	Individuals with a gender identity other than "Man" only (if male sex assigned at birth) or "Woman" only (if female sex assigned at birth) (*ie*, individuals who are not cisgender)
Sexual orientation	Individuals with a sexual orientation other than "Straight" only
Sex assigned at birth	Individuals whose sex assigned at birth is Female, Intersex, or "none of these fully describe me"
Disability status	Individuals who self-report "Yes" to any of the questions about disability status
Disadvantaged background	Individuals who self-report "Yes" to "Are you an individual from a disadvantaged background, as defined by NIH Diversity in Extramural Programs?"

This information is collected by the *All of Us* Research Program through the researcher registration process, during which researchers self-report their demographic information. Responding to the demographic survey is not required, and researchers have the option to select more than one response, "prefer not to answer," and "none of these fully describe me." This information is reported externally only in deidentified counts of more than 20.

Increasing researcher diversity is critical to advancing health equity and is aligned with broader NIH strategic planning.[10,11] Because research topic selection can be related to the race and gender of the scientist conducting the research, the lack of UBW researchers studying a dataset has the potential to further limit health equity.[12] National Institutes of Health collected and analyzed data to assess disparities and inclusivity in order to address existing structural inequities in the development of the biomedical workforce.[13–18] Recent articles from the National Academies of Sciences, Engineering, and Medicine and the NIH Advisory Committee to the Director Working Group on Diversity emphasize the importance of training and education opportunities, mentorship, and capacity building in the form of equal access to grants.[19,20] Underrepresented in the biomedical workforce researchers face bias in receiving NIH R01 funding and are cited less in peer-reviewed publications.[21–25] Structural and institutional biases may account for some of these differences; recent studies show that topics often selected by UBW researchers are funded at lower rates because they submit proposals to and are funded by NIH institutes, centers, and offices with lower award rates.[26,27]

Even with awareness and action around systemic inequities, emerging fields of science are not immune to recreating these structures. In 2022, research in five precision medicine research studies across three national consortia identified disparities within their own research teams.[28] Further, some have argued that datasets with diverse participant cohorts are not enough to ensure equitable outcomes in precision medicine and that "upstream inclusion is compromised by downstream exclusion."[29] This article's main objective is to introduce and describe how an adapted social–ecological model can be used to direct key strategies of the REO Branch and its partners use to build a diverse, inclusive, and growing community of researchers using the *All of Us* dataset.

Materials and methods

Overview

All of Us Research Program defined five researcher audiences, identified areas of diversity and inclusion, applied a working definition of researcher diversity to enable the program and its partners to determine their success at engaging multiple researcher communities, created a researcher engagement ecosystem of program partners, and used a social–ecological modeling approach to operationalize researcher engagement activities and assessments.

Researcher audiences

The five audiences were based on career stages identified by the NIH Division of Biomedical Research Workforce[30] and *All of Us*' commitment to making the dataset broadly accessible[1,31]: (1) K-12 students and teachers; (2) undergraduate, graduate, and health care professional students; (3) postdoctoral researchers and early stage investigators; (4) established investigators; and (5) community and citizen scientists, who are defined as laypeople who are trained in research methodology.[32]

Researcher diversity

During the registration process for the Workbench, researchers receive an optional demographic survey that includes questions about racial and/or ethnic identities, disability status, disadvantaged background, sex assigned at birth, sexual orientation, and gender identity.

The *All of Us* researcher engagement ecosystem

The *All of Us* researcher engagement ecosystem—made up of outreach leaders, participants, researchers, and engagement awardees—is modeled after the program's community and participant engagement ecosystem.[33] In June 2024, there were 13 national awardees within the researcher engagement ecosystem, supported through five funding opportunities/announcements (OT-PM-19-003; OT-PM-16-002; OTA-20-009; OT-PM-17-002; OTA-22-006). Three awardees focus exclusively on researcher engagement and outreach, while the other 10 have goals dedicated to researcher as well as community and participant engagement and outreach. Awardee activities were compared to the researcher engagement ecological model to demonstrate how partner activities can be applied to this framework to support the development of a researcher community (Table 2). A more comprehensive list of awardee activities is included in Table S1.

Researcher engagement partners were supported by the national *All of Us* central office, which includes REO program officers, scientific staff, and communications staff. Researcher Engagement and Outreach is one of the four branches within DEO and works with other divisions across *All of Us* to develop, tailor, and implement communications and outreach materials and tools to support knowledge mobilization. The central office also assists in promoting partner events and activities through AllofUs.nih.gov, ResearchAllofUs.org, program newsletters, program social media accounts, and other channels as appropriate.

Adapted social-ecological model of researcher engagement

All of Us Research Program operationalizes researcher engagement with partners through an adapted social–ecological model of researcher engagement (Figure 1). This model, developed and finalized in 2022 and shared with program partners in September 2023, uses a system-based approach to understand the contexts in which researchers are working. The adapted model modifies a public health framework that had been previously applied to numerous health topics.[34] This public health framework was, in turn, based on an earlier ecological systems theory of human development.[6–9] Recently, it has been used to explore community and participant engagement at *All of Us*.[33] This framework has been described in social science strategies to support professional development programs for UBW researchers.[34] For each level of the framework, REO coordinated with partners to develop specific system-driven strategies (Table 2, Figure 1). Included in the "Results" section are descriptions of selected partner activities and their impact on the *All of Us* researcher user base. Although these activities are presented in only one level of the adapted social–ecological model, activities may span multiple levels simultaneously.

Results

As of June 2024, more than 11340 researchers are registered to access the *All of Us* dataset. These researchers represent more than over 800 institutions that have an agreement with the program to access the *All of Us* dataset. About 29% of these researchers self-identify as trainees (Figure 2). As of June 2024, they have created nearly 4000 workspaces within the Researcher Workbench and have authored more than 470 peer-reviewed publications, which are available via PubMed central (Table 3).

Level 1: individual—outreach and awareness

American Association on Health and Disability Outreach for Disability and Health Researchers

One example of the Outreach and Awareness level is American Association on Health and Disability's (AAHD's) work with disability and health researchers. American Association on Health and Disability is a cross-disability national organization that promotes health equity for individuals living with disabilities. In addition to their engagement with community members through a disability consortium of 21 national and local organizations on enrolling in *All of Us*, AAHD fosters awareness and use of the *All of Us* dataset with researchers with disabilities and researchers focusing on the health of persons living with disabilities. As part of their outreach work, AAHD presented two research-focused webinars with a total of 68 registrants and 31 researchers in attendance.

Delta Research and Educational Foundation Research Accelerates Possibilities Sessions

In advancing another strategy, Delta Research and Educational Foundation (DREF) focused on *All of Us* outreach to individual researchers to engage, educate, and provide access to students and faculty at Historically Black Colleges and Universities (HBCUs). In practice, research accelerates possibilities (RAP) sessions are outreach presentations designed around the principles of precision medicine, *All of Us*, and emphasizing the importance of biomedical research. These sessions explore health topics specifically relevant to African American communities and emphasize the significance of involvement in biomedical research, as both *All of Us* research participants and researchers.

Table 2. Process outcomes of the social–ecological model of researcher engagement.

Framework level	Strategy	Level description	Partner examples of operationalization	Outcomes
Individual	Outreach and awareness	Interactions at the level of the individual researcher to increase awareness, such as participation in outreach activities or receipt of materials and information about the program, its data types, and tools available for research	American Association on Health and Disability Outreach for Disability and Health Researchers	Two research-focused webinars with a total reach of 68 researchers
			Delta Research and Educational Foundation Research Accelerates Possibilities (RAP) Sessions	68 RAP sessions that engaged over 3500 individuals; eight HBCUs gained access to the Workbench
Interpersonal	Education and training	Activities that cultivate team science, mentorship, training, and educational support to advance use of the *All of Us* Workbench	Baylor College of Medicine Biomedical Researcher Faculty Summit	64 participating research scholars from 45 institutions
			University of Utah Summer Institute and Curriculum	38 high school biology teachers; 18 community college faculty; two conference workshops
			Xavier University of Louisiana Big Data Training and Seven Star Academy, Inc.	67 researchers and 44 faculty engaged; three conference abstracts produced
Institutional	Capacity building	Engagement of institutions and partners, including institutions with a track record of engaging researchers underrepresented in the biomedical workforce, and community-based organizations to identify needs and assets, build capacity, enhance research activity, and actively utilize program resources	RTI International Researcher Academy Institutional Champions	19 institutions funded across three cohorts, supporting more than 130 researchers
			National Library of Medicine Data and Training and Engagement for Academic Libraries Program	25 participating institutions across two cohorts
			National Alliance for Hispanic Health *Todos Juntos* Researcher Ambassador Network	15 participating institutions; 21 resulting publications; 1554 researchers (14% of all Workbench registered researchers); 24 new research awards
			Asian Health Coalition Increasing Outreach to Researchers	Development of a national researcher engagement strategy leveraging AANHPI-serving organizations
Community	Bridging communities	Activities that catalyze a community of multiple researcher audiences (eg, K-12 students and teachers, undergraduate and graduate students, post-doctoral students and early stage investigators, established investigators, and community and citizen scientists) and participants to find value in the *All of Us* Workbench	Interdisciplinary Guided Network for Investigation, Translation and Equity Data Sandbox and Data Picnic events	90 participating researchers and community members across 12 events
			PRIDEnet Researcher Basecamp	17 early career and community-based researchers participating
Culture/society	Knowledge mobilization	Dissemination of findings through publications, presentations, or curricula development to communities, participants, and researchers to mobilize and translate knowledge	University of Florida Clinical and Translational Science Award/Partners for the Advancement of Community Engaged Research Community Network	Presented at eight conferences Reached 4836 people New researchers registered on the Workbench: 155 Researcher Consultations: 331 Events: 61 New workspaces: 221 National Community Advisory Board meetings: 13 Abstracts/Presentations or papers in progress: 42 Papers submitted: 4 Papers published: 1
			All of Us Researchers Convention	84 abstract submissions and 1200 registrants in 2023 129 abstract submissions and 1790 registrants in 2024

Abbreviation: AANHPI, Asian American, Native Hawaiian, and Pacific Islander.

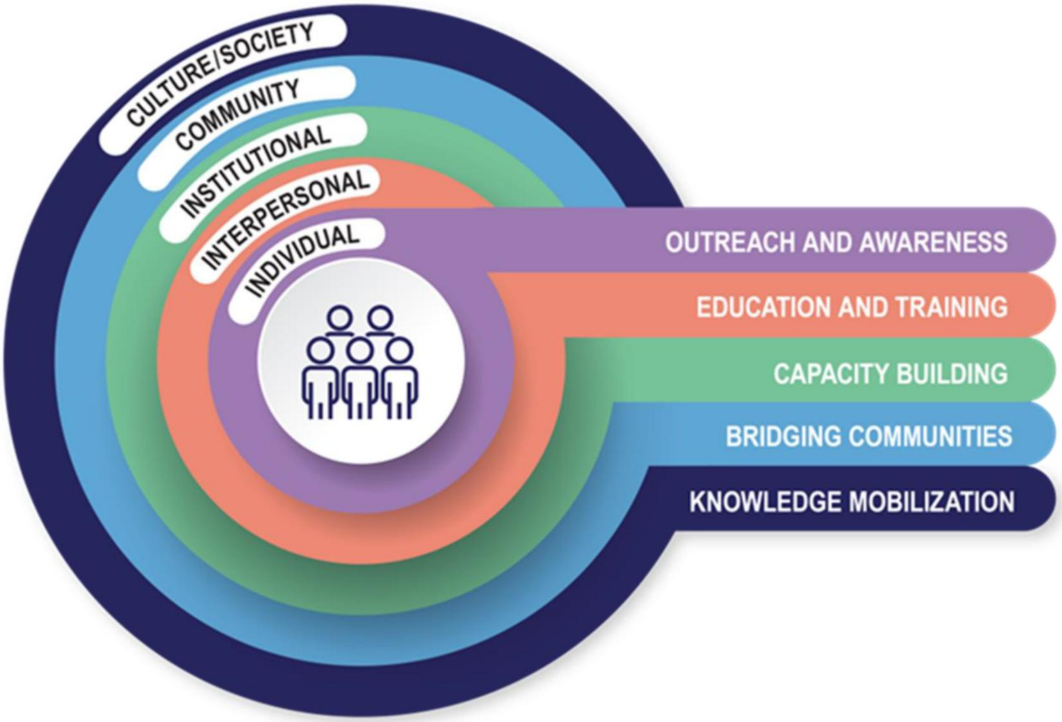

Figure 1. Social–ecological model of researcher engagement.

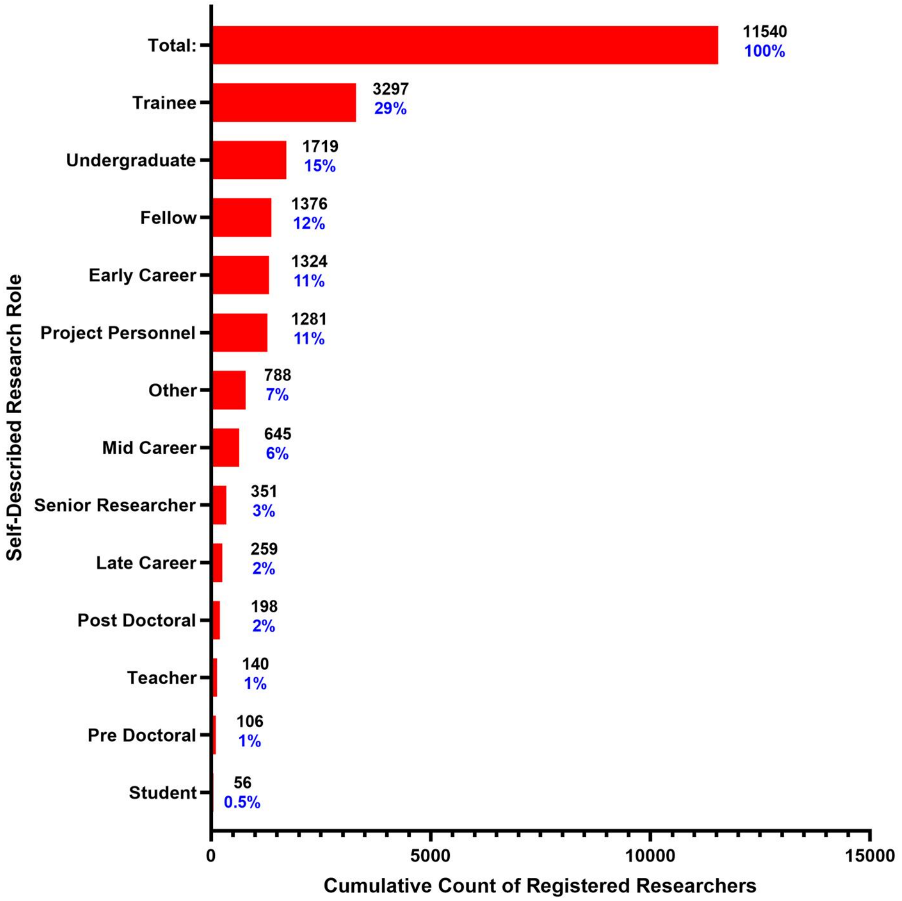

Figure 2. *All of Us* registered researchers by self-described research role.

Table 3. Overview of the *All of Us* researcher community (as of June 2024).

Category	Total	UBW	UBW by SGM	UBW by R/E	UBW by disability	UBW by disadvantaged background
DURAs	812					
Registered researchers	11 347	8913	5864	3656	3499	1805
Workspaces created by active researchers	3997	3176	2032	1355	1316	710
Publications	471	71				

Abbreviations: DURAs, data use and registration agreements; R/E, race and ethnicity; SGM, sexual and gender minority; UBW, underrepresented in the biomedical workforce.

Following the 68 RAP sessions that engaged over 3500 individuals, DREF collaborated with 9 HBCUs, resulting in eight institutions gaining access to the dataset. Delta Research and Educational Foundation Delta Research and Educational Foundation also engaged with 68 researchers who created 36 workspaces on the Workbench.

Level 2: interpersonal—education and training

Baylor College of Medicine, *All of Us* Evenings with Genetics, Biomedical Researchers Faculty Summit

Baylor College of Medicine *All of Us* Evenings with Genetics (BCM AoUEwG) aims to foster interpersonal connections between researchers. From 2022 to 2024, BCM AoUEwG hosted an annual faculty summit to support researchers who are interested in using the Workbench. Scholars work in multidisciplinary teams and participate in a year-long program that offers networking opportunities, the opportunity to apply for external funding, and comprehensive Workbench training. The process of BCM AoUEwG's Workbench training curriculum development is further detailed within this special issue.[35] The first cohort included 30 early career researchers representing 22 institutions, 10 of which are minority-serving institutions (MSIs).[36] The second cohort included 34 researchers from 29 different institutions, 12 of which are MSIs, and the third cohort included 43 researchers from 35 different institutions, including 16 MSIs.

University of Utah Summer Institute and Curriculum

The University of Utah's Genetic Science Learning Center (GSLC) also exemplifies interpersonally driven education and training activities. The GSLC encourages program engagement via education and training materials codeveloped with high school educators for use in high school settings. They partner with the program on the development of materials and course curricula to support awareness and engagement as well as education and training with *All of Us* data. During a 2022 two-and-a-half day summer institute, the GSLC used a method of collaborative curriculum design they developed to co-design course materials for high school educators focused on the publicly available *All of Us* Data Browser.[37,38] At this institute, 19 high school biology teachers developed learning objectives and early concepts for learning activities that the GSLC further developed and validated through classroom pilot testing. The processes for this curriculum development and testing are further detailed within this special issue.[38] Genetic Science Learning Center summer institutes also provided training on using the Workbench for 19 high school biology teachers (2023) and 18 community college faculty (2024). Two conference workshops provided an introduction to the Workbench and the curriculum.

Xavier University of Louisiana Big Data Training and Seven Star Academy, Inc

Xavier University of Louisiana (XULA) developed structured training in R and Python for faculty and student research teams, which translated to more than 700 students learning about *All of Us*. As of April 2024, some 34 faculty members and 10 students have set up workspaces. Xavier University of Louisiana has continued to support the teams through R or Python office hours. Xavier University of Louisiana, an HBCU, also served as an educational resource for other HBCUs beginning in 2021. Since April 2024, 10 HBCU faculty members have registered for the Workbench, set up workspaces, and received R and Python training.

Seven Star Academy, Inc., in collaboration with DREF and XULA, facilitated comprehensive training and ongoing support for researchers, including hands-on workshops and weekly one-on-one research support. During this initiative, 23 researchers from three HBCUs were actively engaged. As a result, three conference abstracts were produced.

Level 3: institutional—capacity building

RTI International *All of Us* Researcher Academy Institutional Champions

To build capacity within institutions, RTI International, in collaboration with Community-Campus Partnerships for Health, developed a comprehensive program to build and maintain a diverse network of researchers using the *All of Us* dataset, primarily within institutions with a track record of supporting UBW researchers and students. One aspect of the Researcher Academy, further detailed in this special issue,[39] is the Institutional Champions (IC) award, which provides up to $75 000 to each awarded institution for 12 months. Research team time (for research or to attend program trainings) is covered via the award, and recipients register for the Workbench, participate in *All of Us* Researcher Academy courses, analyze data, and submit a research abstract to the annual *All of Us* Researchers Convention. Since the establishment of the IC award, 19 institutions have been funded in three cohorts between 2022 and 2024. More than 130 students from multiple institutions have attended Researcher Academy courses. Researchers at the institutions have also submitted four external grant applications, two of which were funded.

National Library of Medicine (NLM) *All of Us* Data Training and Engagement for Academic Libraries Program

The Academic Libraries Program, developed in partnership with National Library of Medicine (NLM), is another example of capacity building for staff who support their campus communities. The Academic Libraries Program is intended to build awareness of and technical expertise for the *All of Us* dataset through training of library staff within academic institutions with a historic commitment to, or track record of, educating UBW researchers. Library staff registered for the Workbench, learned about *All of Us* through tailored trainings and hands-on learning, networked with other program library staff, and created an institutional capacity-building plan. Beginning in 2023, a total of 25 institutions with broad geographic distribution participated in the program, split into two cohorts. Each participating institution received a capacity building award of up to $44 500 for the first cohort and $40 000 for the second, separate cohort. Funds were used for staff time, equipment, engagement activities, and additional training resources for their campus communities.

National Alliance for Hispanic Health Todos Juntos Researcher Ambassador Network

The goal of *Todos Juntos* researcher ambassador network (*Todos Juntos*) was to increase outreach and awareness as well as build capacity by supporting researchers, faculty members, and students as "research ambassadors" within 15 institutions. As of May 2024, there are 1554 registered researchers within the participating institutions (14% of all registered researchers), and the majority (54%) of these researchers self-identify as UBW by race and/or ethnicity. In addition, *Todos Juntos* partnered with Arizona State University to foster research by including topics related to the *All of Us* dataset within the university curriculum as a model for incorporating the Researcher Workbench into educational settings. *Todos Juntos* also supported registered researchers within the 15 participating institutions and created over 750 workspaces, which is about 8% of all workspaces as of June 2024. *Todos Juntos* researchers from eight institutions have also authored 21 publications. In addition, National Alliance for Hispanic Health (NAHH) identified 74 institutions without a DURA and partnered with NIH to support them joining the Researcher Workbench and provided information on the Researcher Workbench to 8884 researchers in professional presentations and conferences. NAHH also partnered with its supporting organization, the Healthy Americas Foundation, to offer Healthy Americas Research Consortium research grants in 2022 and 2023 to 10 teams studying cervical cancer in Hispanic communities. In April 2024, NAHH also issued grants to 14 research teams studying chronic diseases in Hispanic communities. These combined 24 research teams represent a total of 73 researchers expanding the science base for Hispanic health and well-being. In November of 2024, an additional 12 Healthy Americas Research Consortium grants will be made with a focus on Hispanic women's health and the health of Hispanics in rural communities, further expanding this critical network of Hispanic research.

Asian Health Coalition increasing outreach to Asian American, Native Hawaiian, and Pacific Islander researchers in the workbench

In 2018, Asian Health Coalition (AHC) was selected as a community engagement partner to increase the engagement of researchers using the *All of Us* dataset and support capacity building. AHC developed a national engagement strategy that leverages Asian American, Native Hawaiian, and Pacific Islanders (AANHPI)–serving organizations, health care providers, national partners, and researchers to facilitate researcher recruitment.[40,41] AHC also worked with partners to increase understanding of AANHPI data in the Workbench. In 2022, AHC launched the *All of Us* Community Scientist Project (CSP) to invite community-based researchers from AANHPI Engagement and Recruitment Core partner organizations and academic researchers to collaborate on research projects about the AANHPI community in the Workbench. To date, CSP involves 16 researchers from academic institutions and AANHPI-serving organizations. The CSP group collaborated on the conceptualization of research questions, data analysis, and interpretation of the data, which resulted in a group manuscript on AANHPI mental health.

Level 4: community—bridging communities

Interdisciplinary Guided Network for Investigation, Translation, and Equity Data Sandbox and Data Picnic events

The fourth level of the model focuses on connections between an institution and the communities within which it exists. Feinstein Institutes for Medical Research of Northwell Health aims to increase researcher access and impact and simultaneously return value to communities through Data Sandbox and Data Picnic events. The Data Sandbox is a workshop about the public *All of Us* Data Browser for between 15 and 50 community-based researchers with shared occupations or interests. The Data Picnic is a workshop series for groups of fewer than 10 academics and researchers that introduces the Workbench in 90-minute sessions. Since 2022, Interdisciplinary Guided Network for Investigation, Translation, and Equity (IGNITE) has hosted 12 of these events with over 90 researchers and is facilitating the community-to-researcher relationship through similar activities in 2024.

PRIDEnet Researcher Basecamp

Stanford University's PRIDEnet has undertaken activities to build capacity with academic and community researchers to understand and conduct meaningful research with a special focus on research topics related to sexual and gender minority populations. In December 2020, PRIDEnet hosted a 3-session journal club to increase community members' skills and comfort with research. That journal club led to three additional training activities to support researchers' use of the *All of Us* dataset. In summer 2022, PRIDEnet recruited four undergraduate summer interns, teaching them about community-engaged research and about *All of Us*. In March 2023, PRIDEnet then hosted an in-person, two-day Researcher Basecamp to build data science skills for 17 early career and community-based researchers, resulting in 20 newly registered researchers on the Workbench.

Level 5: culture/society—knowledge mobilization

University of Florida Clinical and Translational Science Award/Partners for the Advancement of Community Engaged Research Community Network

The fifth level of the adapted model connects to the contexts within which communities themselves exist. These strategies emphasize connecting societal impacts. A key example of this is the Clinical and Translational Science Award/Partners for the Advancement of Community Engaged Research Community Network (CPCN), which is supported by Health Literacy Media, a subawardee of the University of Florida. CPCN engages community-engaged clinical and translational science researchers from eight institutions across the United States. Since 2023, CPCN has presented on *All of Us* topics at 61 events, including eight conferences. As a result of these efforts, more than 330 researchers registered for the Workbench and created more than 200 workspaces on the platform. Nearly all of the more than 60 Clinical and Translational Science Awards attend the monthly Partners for the Advancement of Community Engaged Research meetings and hear about the *All of Us* program with encouragement to use the Workbench. The CPCN also collaborates with community partners in determining research topics for researchers to explore on the Workbench and community-engaged research dissemination strategies and encourages community and academic partner coauthorship of publications. As of March 2024, members of the CPCN have produced more than 40 abstracts, posters, presentations, and manuscripts using *All of Us* data.

All of Us Researchers Convention

The *All of Us* Researchers Convention, managed by Pyxis Partners, exemplifies knowledge mobilization. The convention is a virtual assembly of researchers, participants, and the broader community that aims to share research conducted with *All of Us* data and recognize the essential role of participants in the research process. First held in May 2021, the convention was originally developed by the Black Greek Letter Consortium and the Southern Network, a Community and Provider Gateway Initiative partner. More than 1200 individuals registered to attend with 596 attending the second annual convention held in March 2023. Presenters are selected from researcher and student researcher pools. For the researcher pool, more than 80 abstracts were submitted, and 15 researchers were selected to present. More than 20% of total submissions were from MSIs. For the student researcher pool, student researchers from the *All of Us* Research Scholar Program and *All of Us*–affiliated community organizations were also invited to participate in the event through a digital poster session. Twelve students, including one high school student, were selected to present their research using *All of Us* data out of a total of 58 poster submissions. More than 1700 people registered for the 2024 convention with 862 attending.

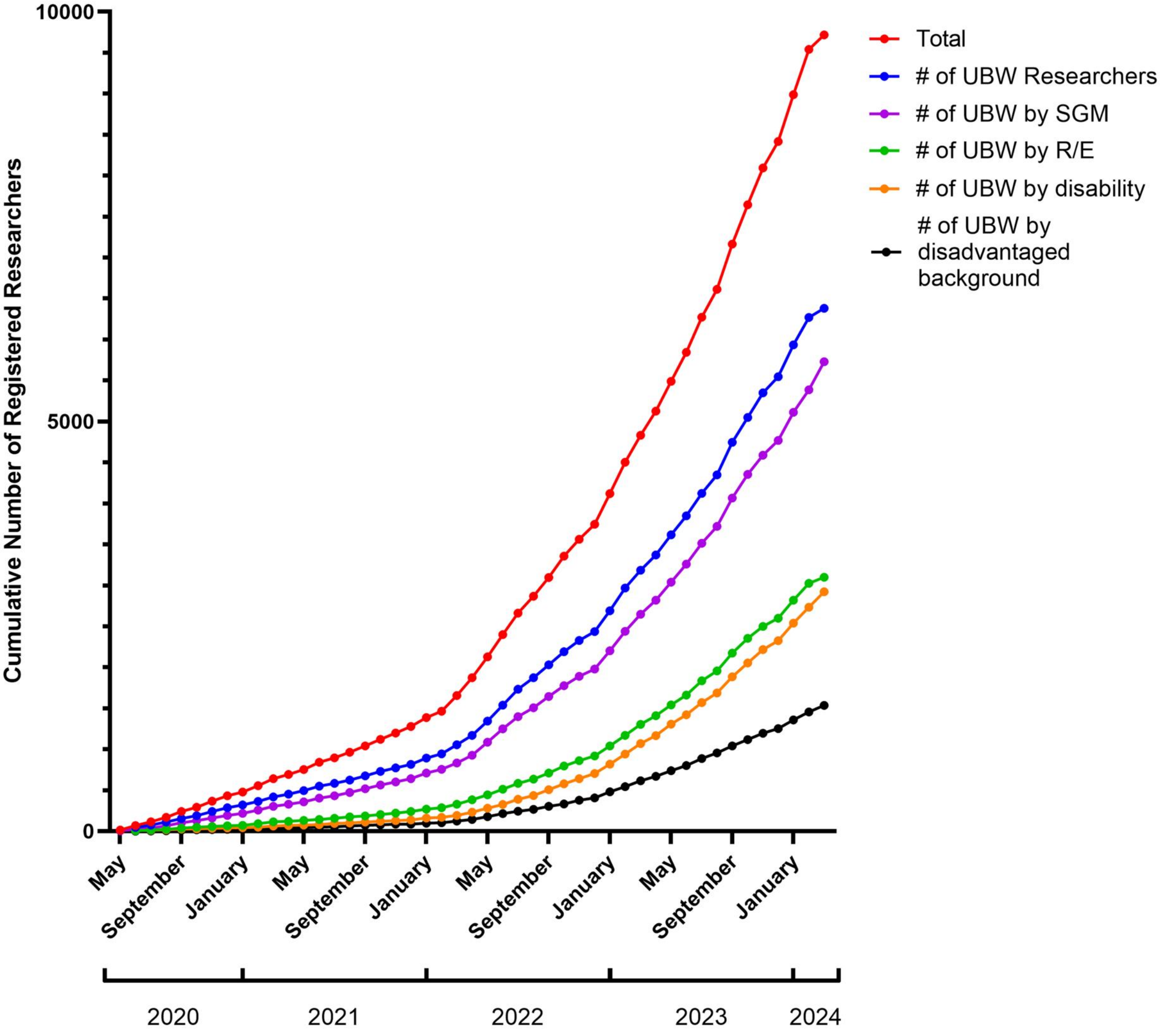

Figure 3. *All of Us* researcher workbench user count by demographic categories underrepresented in the biomedical workforce (UBW), May 2020–March 2024. Abbreviations: R/E, racial and/or ethnic minority; SGM, sexual and/or gender minority.

Results—*All of Us* researcher community growth and diversity

As a result of the intentional engagement of our researcher ecosystem partners, the *All of Us* researcher community is diverse by researcher demographics as well as institution type. As of March 2024, about 66% of *All of Us* registered researchers are considered UBW (Figure 3). About 59% of UBW researchers overall are from sexual and gender minority groups (Figure 3). About 36% of *All of Us* researchers self-identify as members of racial and ethnic minority groups and about 35% self-identify as a person living with a disability. Just over 18% of *All of Us* researchers self-identify as individuals from disadvantaged backgrounds (as defined by NIH; NOT-OD-20-031).[27] In 2023, UBW researchers accounted for two-thirds of all new registered users (67%, Figure 4). About 1 out of every 5 institutions with access to the dataset is an MSI (Figure 5). As of June 2024, 40% of HBCUs had access to the data (including all four HBCU medical schools), and 25% of all AANHPI and 12% of all HSIs also had access. Notably, 67% researchers who have authored a manuscript using *All of Us* data self-identify as UBW (Figure S2). Finally, about 15% of published researchers are affiliated with MSIs (Figure S3).

Discussion

All of Us Research Program has established a diverse, inclusive, and growing research community by building capacity among diverse researcher audiences and by creating an ecosystem of partners guided by an ecological model of researcher engagement. Similar system-based strategies have been identified by other researchers as critical for building a diverse biomedical workforce.[42] The development of this broad researcher community could also have broad health impacts—previous studies indicate that addressing disparities in a biomedical workforce can support health equity.[43–45] Of note, the percentage of all UBW subcategories increased year over year for 2020–2023, while the total percentage of new UBW researchers stayed consistent during this time. This could indicate that newly registered researchers are more likely to identify as members of multiple UBW groups (Figure 4). The increase in the total number of registered researchers could also be partially attributed to changes to the analytics platform itself. In March 2022, the program launched the Controlled Tier, which included genomic data from samples contributed by *All of Us* participants. This likely caused a short-term spike in registrations and workspace creations in 2022.

The program also reaches out to researchers who self-identify as American Indian and Alaskan Native (AI/AN). The Tribal Engagement and Outreach Branch within the *All of Us* Division of Engagement and Outreach leads efforts to engage Tribal communities and AI/AN researchers. Early engagement activities with Tribal researchers include a partnership with the American Indian Science and Engineering Society (AISES) and a coding workshop at the 2003 Annual AISES Conference. The Indigenous Research Working Group utilizes Indigenous scientific knowledge from research professionals to support responsible data use. Additional partnerships, including Tribal colleges and universities, hold great

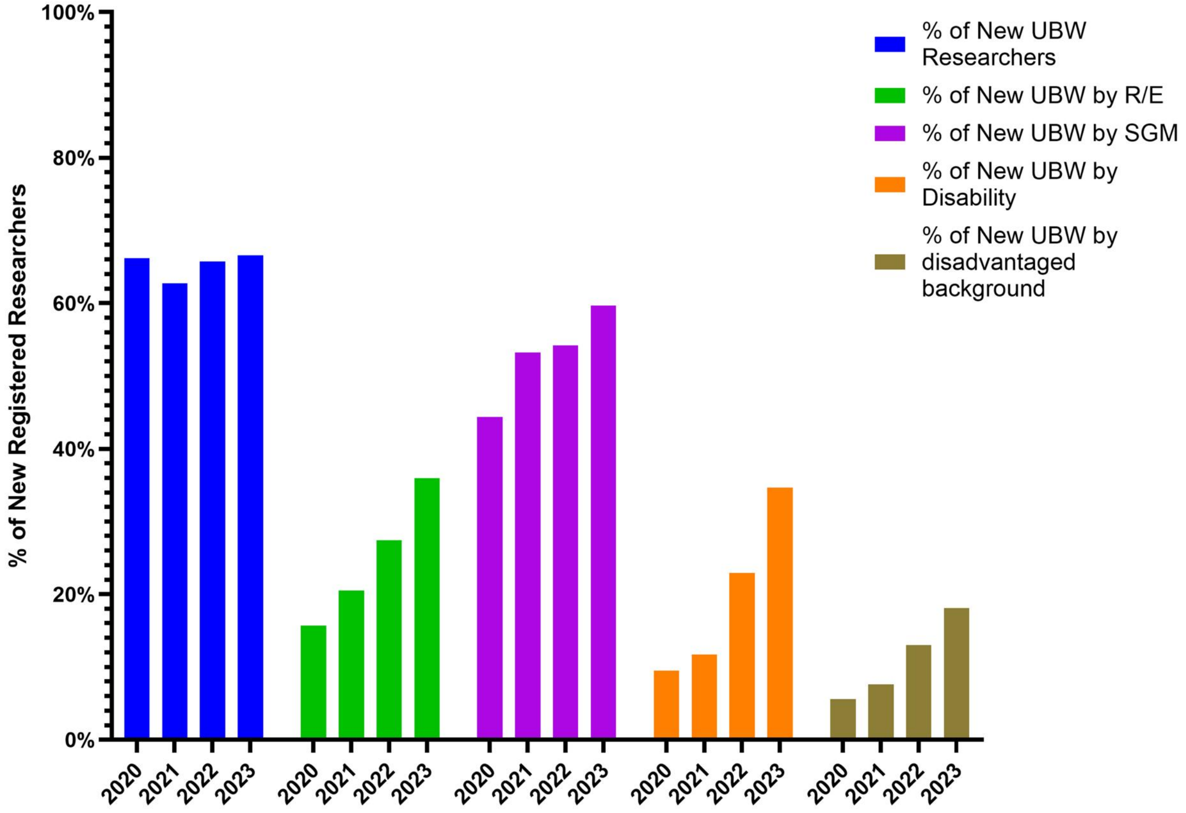

Figure 4. *All of Us* researcher workbench new users by demographic categories underrepresented in the biomedical workforce (UBW), May 2020–March 2024. Abbreviations: R/E, racial and/or ethnic minority; SGM, sexual and/or gender minority.

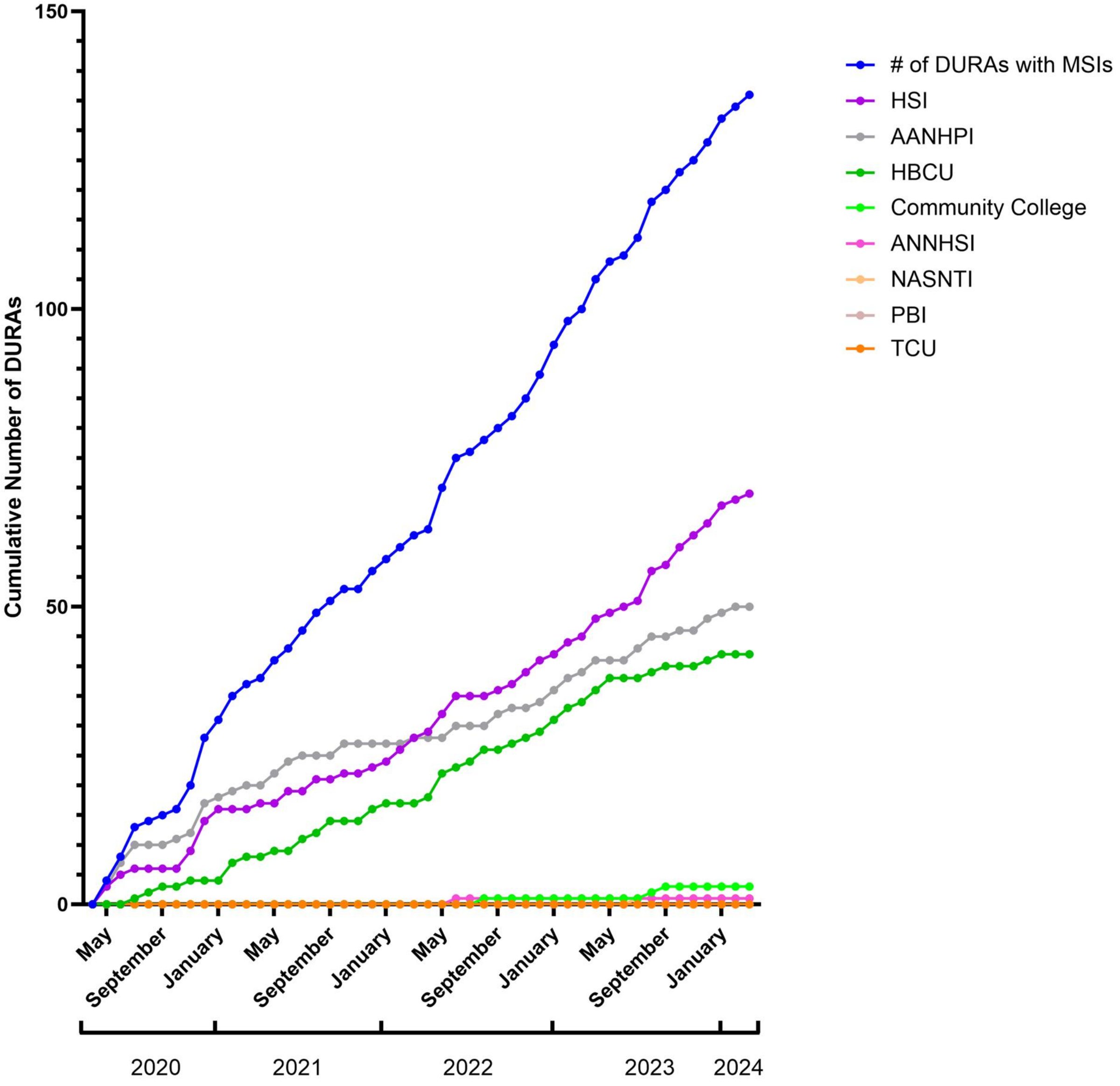

Figure 5. Institutions signing a Data Use and Registration Agreement (DURA) with *All of Us*, May 2020–March 2024. The total number of DURAs has increased at an approximately linear rate from 26 in May 2020 to 664 in March 2024. Abbreviations: AANHPI, Asian American- and Native Hawaiian/Pacific Islander-serving institutions; ANNH, Alaska Native- and Native Hawaiian-serving institutions; HBCU, Historically Black Colleges and Universities; HSI, Hispanic-serving institutions; MSI, minority-serving institutions; NASNTI, Native American-serving nontribal institutions; PBI, Predominantly Black institutions; TCU, Tribal colleges and universities. All analyses of MSIs were conducted post-DURA signing.

potential for equitable research and professional development.

While there have been programmatic successes and benefits of the *All of Us* dataset, several challenges exist as barriers to building capacity among our research community. Until April 2024, researchers had to be proficient in R and Python to work in the Workbench. This created barriers for researchers who do not know these programming languages. In March 2024, the program made RStudio available on the Workbench. This program is more familiar to some researchers than Jupyter Notebook, which had been available since launch. In April 2024, the program released SAS Studio on the Workbench. This provided the first opportunity to use a programming language other than R or Python.

Additionally, while access to the Workbench is free, researchers must pay for cloud computing costs. Researchers receive $300 in initial credits. However, depending on the type of analysis, research teams may quickly run out of initial funds. As noted in the introduction to this article, this could pose particular challenges for UBW researchers and MSIs that are funded at lower rates. These challenges could be partly addressed by several existing NIH resources, such as enrolling in the NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability Initiative and using the NIH Cloud Lab for estimating computing costs and testing analytic code. However, additional resources in the form of trainings and supplementary funding may be warranted.

Despite these challenges, the *All of Us* approach could be applied to other large data repositories interested in building an inclusive user base. The Science Collaborative for Health disparities and Artificial intelligence bias REduction, NHLBI BioData Catalyst, NHGRI Analysis Visualization and Informatics Lab-space, Trans-Omics for Precision Medicine, and many other research programs have all built active researcher communities. The approaches discussed above could be applied to the development of researcher communities within or outside the United States.

Conclusion

Inequities within the biomedical research field are rooted in systemic issues and demand a comprehensive system-based response. Our adaptation of an ecological model generates insights into the systems within which researchers exist and into conditions that support all researchers. This approach considers future researcher engagement in terms of not only pipelines or pathways but also environments.

Acknowledgments

The authors would like to note the support to researchers using the *All of Us* dataset provided by the *All of Us* Data and Research Center. The authors would also like to thank Drs Lilanta Bradley and Stacy Jones for developing the initial concept for the Researchers Convention. Additionally, the authors would like to acknowledge Sheila James for her work as the Program Officer for the Asian Health Coalition and Al Richmond for his work with the *All of Us* Research Program as the Executive Director of Community-Campus Partnerships for Health. The authors gratefully acknowledge *All of Us* participants for their contributions, without whom this research and this research community would not have been possible.

Author contributions

This work was carried out in part by staff of the National Library of Medicine (NLM), National Institutes of Health, with support from NLM.

Rubin Baskir provided conceptualization, formal analysis, data visualization, and led manuscript writing in his role as Director for the Researcher Engagement and Outreach Branch in the Division of Engagement and Outreach for the *All of Us* Research Program and Program Official for RTI International and University of Utah. Minnkyong Lee provided conceptualization support as well as review and editing as Deputy Director of the Division of Engagement and Outreach and as the *All of Us* liaison for the National Library of Medicine. Sydney McMaster provided conceptualization, data curation, visualization of the social-ecological model of researcher engagement and provided review and editing, as well as providing original writing as a Program Official for Baylor College of Medicine, Stanford University PRIDEnet, and University of Florida. Jessica Lee provided visualization, review, and editing support for the manuscript. Shanté Alerté provided review and editing support for the manuscript. Yong Crosby, Faith Blackburne-Proctor, Erica Zumba, Michael Hahn, and Ahmed Elmi provided writing, review, and editing support as Program Officials for the American Association on Health and Disability, Delta Research and Educational Foundation, Feinstein Institute for Medical Research, the National Alliance for Hispanic Health, University of Arizona, and Pyxis Partners, respectively. Principal Investigators Debra Murray and Brendan Lee, Cheryl Thomas, Mitchell Lunn and Juno Obedin-Maliver, Linda Cotler and Milton "Mickey" Eder, Fornessa Randall, Megan Lewis and Jennifer Uhrig, Laura Bartlett, Edgar Gil Rico and Adolph Falcón, Elizabeth Cohn, Jason Karnes, Ronnie Tepp, and Louisa Stark provided project administration, writing, review and editing to the manuscript in their roles as awardees within the researcher engagement ecosystem. Sula Hood provided review and editing to the manuscript and served an integral role in the development of the Researcher Academy. KiTani Lemieux provided program administration and writing review and editing for Xavier University of Louisiana activities. Nelson Lemieux III and Nelson Lemieux Jr. provided program administration and writing review and editing for Seven Star Academy, Inc. activities. Romuladus Azuine and Nakia Mack provided project administration for the Data and Research Center (DRC) and reviewed and edited the manuscript. Martin Mendoza and Janeth Sanchez provided review and editing for the manuscript in their respective roles as Health Equity Director and Deputy Director for *All of Us*. Elizabeth Stewart provided data curation for researcher data as a member of the DRC and Danielle Wilfong provided review and editing for DRC projects within the manuscript. Meag Doherty provided review and editing from her role as a member of the user experience team within *All of Us*. Margaret Farrell and Grace Hébert provided review and editing support as members of the *All of Us* Division of Communications and Marketing. Lilanta Bradely provided conceptualization for the Researchers Convention. Monica Rodriguez and Meredith Wilson provided program administration and writing review and editing as Pyxis Partner team members facilitating the Researchers Convention. Chris Lunt provided writing review and editing in his role as the *All of Us* Chief Technology Officer. Finally, Karriem Watson provided conceptualization, supervision, original draft writing, review, and editing in his role as the Director of the Division of Engagement and Outreach.

Supplementary material

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

Funding

The *All of Us* Research Program and this work is supported by the National Institutes of Health, Office of the Director: Regional Medical Centers: 1 OT2OD026549; 1 OT2OD026554; 1 OT2OD026557; 1 OT2OD026556; 1 OT2OD026550; 1 OT2OD 026552; 1 OT2OD026553; 1 OT2OD026548; 1OT2OD026551; 1 OT2OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085U; Data and Research Center: 5 U2COD023196; Biobank: 1 U24OD023121; The Participant Center: U24OD023176; Participant Technology Systems Center: 1U24 OD023163; Communications and Engagement: 3 OT2OD023205; 3 OT2OD023206; and Community Partners: 1 OT2OD025277; 3 OT2OD025315; 1 OT2OD025337; 1 OT2OD025276; 3 OT2OD031919.

Conflicts of interest

J.O.-M. has consulted for Sage Therapeutics (2017), Ibis Reproductive Health (2017-2018, 2020-present), Hims, Incorporated (2019-present), Folx, Incorporated (2019-present), and Upstream Inc. (2024). None of these engagements influenced or are pertinent to the work described in this manuscript. All other authors have no competing interests to declare.

Data availability

Data cannot be shared outside of the program for ethical/privacy reasons.

References

1. Denny JC, Rutter JL, Goldstein DB, et al.; All of Us Research Program I. The "All of Us" Research Program. *N Engl J Med*. 2019;381:668-676.
2. Mayo KR, Basford MA, Carroll RJ, et al. The All of Us Data and Research Center: creating a secure, scalable, and sustainable ecosystem for biomedical research. *Annu Rev Biomed Data Sci*. 2023;6:443-464.
3. National Institutes of Health (NIH). Notice of NIH's Interest in Diversity, Notice Number: NOT-OD-20-031. NIH, Division of Biomedical Research Workforce, Office of Extramural Research. Updated November 22, 2019. Accessed April 2, 2024. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-20-031.html
4. Asian Health Coalition. Asian Health Coalition, improving health & wellness for Asian, Pacific Islander, African, and other communities of color. AHC. Updated 2023. Accessed April 2, 2024. https://www.asianhealth.org/
5. Obedin-Maliver J, Hunt C, Flentje A, et al. Engaging sexual and gender minority (SGM) communities for health research: building and sustaining PRIDEnet. *J Community Engagem Scholarsh*. 2024;16:9.
6. Bronfenbrenner U. Toward an experimental ecology of human development. *Am Psychol*. 1977;32:513-531.
7. McLeroy K, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. *Health Educ Q*. 1988;15:351-377.
8. Krug E, Dahlberg T, Mercy J, et al. *World Report on Violence and Health*. World Health Organization; 2002.
9. Dahlberg L, Krug E. *Violence: A Global Public Health Problem*. World Health Organization; 2022.
10. Fiscal Years 2023-2027. NIH-wide strategic plan for diversity, equity, inclusion, and accessibility. Updated 2023. Accessed April 6, 2024. https://www.nih.gov/sites/default/files/about-nih/nih-wide-strategic-plan-deia-fy23-27.pdf
11. UNITE Progress Report: 2021-2022. Updated 2021. Accessed April 6, 2024. https://www.nih.gov/sites/default/files/research-training/initiatives/ending-structural-racism/UNITE-progress-report-2022.pdf
12. Kozlowski D, Lariviere V, Sugimoto CR, Monroe-White T. Intersectional inequalities in science. *Proc Natl Acad Sci U S A*. 2022;119:2.
13. Valantine H. NIH's essential 21st-century research challenge: enhancing scientific workforce diversity. *J Invest Dermatol*. 2016;136:2327-2329.
14. Valantine HA. Where are we in bridging the gender leadership gap in academic medicine? *Acad Med*. 2020;95:1475-1476.
15. Valantine HA. NIH's scientific approach to inclusive excellence. *FASEB J*. 2020;34:13085-13090.
16. Valantine HA, Collins F. NIH progress toward inclusive excellence. *Science*. 2020;367:1204.
17. Valantine HA, Collins FS. National Institutes of Health addresses the science of diversity. *Proc Natl Acad Sci U S A*. 2015;112:12240-12242.
18. Valantine HA, Lund PK, Gammie AE. From the NIH: a systems approach to increasing the diversity of the biomedical research workforce. *CBE Life Sci Educ*. 2016;15:fe4.
19. Advisory Committee to the Director Working Group on Diversity. *Racism in Science Report*. National Institutes of Health; 2021.
20. Vargas EA, Scherer LA, Fiske ST, Barabino GA, eds. Minoritized individuals in STEMM: consequences and responses to racial bias and how STEMM professionals can help. In: *Advancing Antiracism, Diversity, Equity, and Inclusion in STEMM Organizations: Beyond Broadening Participation*. National Academies Press; 2023:117-148.
21. Ginther DK. Reflections on race, ethnicity, and NIH research awards. *Mol Biol Cell*. 2022;33:ae1.
22. Ginther DK, Basner J, Jensen U, et al. Publications as predictors of racial and ethnic differences in NIH research awards. *PLoS One*. 2018;13:e0205929.
23. Ginther DK, Kahn S, Schaffer WT. Gender, race/ethnicity, and National Institutes of Health R01 research awards: is there evidence of a double bind for women of color? *Acad Med*. 2016;91:1098-1107.
24. Ginther DK, Schaffer WT, Schnell J, et al. Race, ethnicity, and NIH research awards. *Science*. 2011;333:1015-1019.
25. National Institutes of Health (NIH). Trends in race/ethnicity of NIH-supported Ph.D. recipients. Updated November 2023. Accessed April 2, 2024. https://report.nih.gov/nihdatabook/report/270
26. Hoppe TA, Litovitz A, Willis KA, et al. Topic choice contributes to the lower rate of NIH awards to African-American/Black scientists. *Sci Adv*. 2019;5:eaaw7238.
27. Lauer MS, Doyle J, Wang J, Roychowdhury D. Associations of topic-specific peer review outcomes and institute and center award rates with funding disparities at the National Institutes of Health. *eLife*. 2021;10
28. Jeske M, Vasquez E, Fullerton SM, et al. Beyond inclusion: enacting team equity in precision medicine research. *PLoS One*. 2022;17:e0263750.
29. Galasso I. Precision medicine for whom? Public health outputs from "Genomics England" and "All of Us" to make up for upstream and downstream exclusion. *Am J Bioeth*. 2024;24:71-85.
30. National Institutes of Health Research Training and Career Development (NIH RTCD). Career path. Accessed April 2, 2024. https://researchtraining.nih.gov/career-path
31. Precision Medicine Initiative (PMI) Working Group. *Report to the Advisory Committee to the Director, NIH*. PMI Working Group; 2015.
32. Watson KS, Siegel LD, Henderson VA, et al. The SHARED Project: a novel approach to engaging African American men to address lung cancer disparities. *Am J Mens Health*. 2020;14:1-11.
33. Cohn E, Tepp R. Promising practices for inclusive precision medicine research and the contribution to public and population health. *Am J Bioeth*. 2024;24:1-4.
34. Matthews AK, Allen-Meares P, Watson K, et al. The use of strategies from the social sciences to inform pipeline development programs for under-represented minority faculty and students in the health sciences. *J Clin Transl Sci*. 2020;5:e73.
35. Coleman JR, Baker JN, Ketkar S, et al. Data science training curriculum for the All of Us Researcher Workbench. *JAMIA*. 1-12. Under Review.
36. Higher Education Act (HEA), 20 U.S.C. § 1067k(3). Accessed October 25, 2024. https://www2.ed.gov/about/offices/list/ocr/edlite-minorityinst-list.html
37. Drits-Esser D, Stark LA. The impact of collaborative curriculum design on teacher professional learning. *Electron J Sci Educ*. 2015;19:1-27.

38. Stark LA, Fenker KE, Krishnan H, et al. Research to classrooms: a co-designed curriculum brings All of Us data to secondary schools. *JAMIA*. 2024:ocae167. https://doi.org/10.1093/jamia/ocae167
39. Southwell B, Hood S, Carter J, et al. A model for supporting biomedical and public health researcher use of publicly available All of Us data at Historically Black Colleges and Universities. *J Am Med Inform Assoc*. 2024:1-5. https://doi.org/10.1093/jamia/ocae099.
40. Randal FT, Lozano P, Qi S, et al. Achieving a representative sample of Asian Americans in biomedical research through community-based approaches: comparing demographic data in the All of Us Research Program with the American Community Survey. *J Transcult Nurs*. 2023;34:59-67.
41. Randal FT, Qi S, Lozano P, et al. Establishing a national engagement strategy for recruiting Asian Americans and other minorities into biomedical research. *Prog Community Health Partnersh*. 2022;16:349-359.
42. Genomic Data Science Community Network (GDSCN). Diversifying the genomic data science research community. *Genome Res*. 2022;32:1231-1241.
43. Swartz TH, Palermo AS, Masur SK, et al. The science and value of diversity: closing the gaps in our understanding of inclusion and diversity. *J Infect Dis*. 2019;220:S33-S41. https://doi.org/10.1093/infdis/jiz174
44. Winkfield KM, Levit LA, Tibbits M, et al. Addressing equity, diversity, and inclusion of Black physicians in the oncology workforce. *JCO Oncol Pract*. 2021;17:224-226. https://doi.org/10.1200/OP.21.00079
45. Carnethon MR, Kershaw KN, Kandula NR. Disparities research, disparities researchers, and health equity. *JAMA*. 2020;323:211-212.

Published by Oxford University Press on behalf of the American Medical Informatics Association 2024.

Journal of the American Medical Informatics Association, 2024, 32, 38–50
https://doi.org/10.1093/jamia/ocae270
Research and Applications