

# Community Engagement Workshop

*National Institutes of Health*

*Office of Strategic Coordination – NIH Common Fund*

July 24-25, 2024

Virtual Meeting

*Revised on October 16, 2024*



This meeting summary was prepared by Rose Li and Associates, Inc., under contract to the National Institutes of Health's (NIH) Office of Strategic Coordination (OSC). The views expressed in this document reflect both individual and collective opinions of the meeting participants and not necessarily those of NIH OSC. Contributions to this summary by the following individuals are gratefully acknowledged: Joshua Elmore, Juliana Piacentini, Erica Schipper, Bethany Stokes, Jessica Tennis, and Nancy Tuvesson.

## Table of Contents

**Acronym List ..... iii**

**Meeting Summary ..... 1**

Welcome and Introduction.....1

Keynote Presentation: Revisiting Preconceived Notions of Community Engagement .....1

NIH Common Fund and Common Fund Data Ecosystem Resources.....3

Panel Discussion: How Data Policy (Privacy and Ethics) Impacts Historically Underserved  
Communities .....4

Workshop Breakout Session: Community Engagement Strategies .....6

Panel Discussion: Data Use and Accessibility.....8

Workshop Breakout Session: Data Challenges and Barriers to Data Use .....10

Panel Discussion: Training & Education Needs of the Data Community.....11

Workshop Breakout Session: Data Solutions to Identified Barriers from Day 1.....14

Workshop Breakout Session: Immediate Training Needs and Supports of the Community .....16

Workshop Closing Presentation & Wrap Up.....17

**Appendix A: Agenda ..... 19**

**Appendix B: Meeting Speakers and Personnel ..... 21**

**Appendix C: Workshop Breakout Sessions Discussion Questions ..... 23**

**Appendix D: Website Links ..... 25**

## Acronym List

AI/AN	American Indian and Alaskan Native
CARE	Collective Benefit, Authority to Control, Responsibility, and Ethics
CBO	community-based organization
CBPR	Community-Based Participatory Research
CDR	Catalytic Data Resources
CFDE	Common Fund Data Ecosystem
DCC	Data Coordinating Center
FAIR	findable, accessible, interoperable, reusable
HINTS	Health Information National Trends Survey
IRB	Institutional Review Board
LGBTQIA+	lesbian, gay, transgender, queer, intersex, and asexual
NCI	National Cancer Institute
NIH	National Institutes of Health
NLM	National Library of Medicine
ODSS	Office of Data Science Strategy
OSC	Office of Strategic Coordination

## Meeting Summary

### Day 1: July 24, 2024

#### Welcome and Introduction

*Douglas Sheeley, Sc.D., OSC, NIH*

*Natalie Vineyard, M.S., OSC, NIH*

Dr. Sheeley, Acting Director of the National Institutes of Health (NIH) Office of Strategic Coordination (OSC), welcomed attendees to the NIH Common Fund Community Engagement Workshop. The workshop aimed to identify ways to improve the utility and interoperability of NIH Common Fund data resources and to build relationships with representatives from underrepresented communities in biomedical research. The workshop also sought to explore effective outreach approaches and develop strategies to help these communities utilize Common Fund resources more effectively.

#### Keynote Presentation: Revisiting Preconceived Notions of Community Engagement

*Krystal Tsosie, Ph.D., M.P.H., M.A., Arizona State University*

Community engagement is a constantly evolving process that should be rooted in community-driven innovation and decision-making processes. However, many current researchers continue to engage in “engagement washing”—a process in which researchers recognize the ethical imperative to engage with stakeholders but pursue such engagement only to satisfy reputational aims. Researchers should instead aim to integrate community involvement *throughout* the research process, from the design of research questions to the dissemination of findings. This model, known as Community-Based Participatory Research (CBPR), has been applied to various fields, including environmental justice research since the 1970s, but remains relatively new to many biomedical fields. Tribally-Driven Research, an evolution beyond CBPR, emphasizes community ownership and decision-making at every stage of research for American Indian and Alaskan Native (AI/AN) populations.

Theoretically, community engagement should involve thorough research, qualitative training, and strong local infrastructure. However, in practice, effective engagement can be obstructed by the lack of researcher training or local infrastructure, difficulties in identifying key local experts, conflicting policies between institutions and communities, and the misalignment of academic timelines with the needs of communities. Engagement in research involving AI/AN individuals is further complicated by decades of assimilation policies and the economic displacement of tribal members, resulting in 72% of AI/AN individuals now residing in urban areas. To address these challenges, sustainable policies must be developed to ensure true community engagement and respect for tribal sovereignty in research. Initiatives, such as Indigenous Data Science ([IndigiData](#)), propose a promising approach to promote training among native populations. IndigiData focuses on indigenous-led data infrastructures. This program has

trained more than 60 indigenous students in informatics, coding, data ethics, indigenous data sovereignty, and cultural pedagogies. By involving tribal communities and hosting workshops on tribal lands, IndigiData contributes to local research training capacity, reinvests in communities, and views data as a critical resource for sustainable development that will benefit tribal communities going forward.

Indigenous communities often face pressure to participate in research under the assumption that failing to do so will result in missing out on potential benefits. This assertion fails to acknowledge the historical power imbalances that have led to health disparities and their mistrust of the research community. Research that uses indigenous genomic data rarely benefits these communities, because companies consider this work to be not profit-generative because of small population sizes. The [Native BioData Consortium](#), founded in 2018, seeks to address these inequities by creating a data trust that prioritizes the needs and perspectives of indigenous communities. This consortium uses a dynamic consent portal, blockchain technology, and federated approaches to ensure that indigenous peoples have control over their data and can prevent unauthorized secondary use.

Early in the COVID-19 pandemic, NIH asked tribal nations to participate in both [Rapid Acceleration of Diagnostics \(RADx\)](#) initiative and the [All of Us](#) Research Program to collect COVID-19-related data. The National Congress of American Indians expressed concerns about the equity of these programs, noting that existing NIH awardees' chances for support appear to be better than those of new applicants. It also sought clarification on how NIH would ensure respectful partnerships with tribal nations and address data sharing and intellectual property rights. This criticism reflects broader concerns among tribal nations and their community members about the collection and use of their data, including questions about who benefits from this research. *All of Us*, which focuses on recruiting underrepresented groups for research, started to recruit indigenous individuals in 2018, but only started to consult with tribal nations in 2019. This delay led tribal leaders and reporters to accuse NIH of bypassing tribal sovereignty to harvest genomic data from AI/AN communities. Early program language equated consultation with engagement, which indigenous scientists criticized as a unidirectional approach. *All of Us* has since established a tribal working group and new engagement policies to address these issues, emphasizing the need to equitably train scholars from indigenous communities.

In response to community concerns, the RADx launched a Tribal Data Repository. This repository, created in collaboration with tribal nations, is designed to house tribal data that tribal communities can trust. It aims not only to build data and policy infrastructure, but also to establish research capacity and education resources for AI/AN scholars, tribal government entities, and tribal communities. The repository represents a significant step toward integrating tribal data governance within existing federal and international policies and creating a research and analytic framework that empowers tribal nations to access and benefit from their own data.

## NIH Common Fund and Common Fund Data Ecosystem Resources

*George Papanicolaou, Ph.D., OSC, NIH*

### Overview

[The Common Fund](#) is an NIH funding entity that seeks to accelerate emerging science, remove research roadblocks, enhance the research workforce, and support high-risk, high-reward science. The Common Fund supports high-impact, trans-NIH programs that address ambitious goals in biomedical research. These programs must be transformative in nature, short term and goal driven, cross-cutting, synergistic, and novel. They are designed to produce deliverables within 10 years or less to address complex issues requiring trans-NIH teams and provide new solutions to specific challenges. The management of these programs involves collaboration across diverse scientific disciplines and NIH institutes, ensuring coordination and preventing duplication of efforts.

The Common Fund supports three types of programs: (1) Transformational Science and Discovery programs, which aim to establish new scientific principles and models; (2) Catalytic Data Resources (CDR) programs, which focus on managing and developing data for scientific discoveries; and (3) Re-Engineering the Research Enterprise programs, which are designed to transform the conduct of biomedical and behavioral research. Each of these types of programs plays a role in advancing the overall mission of the Common Fund.

### The Common Fund Data Ecosystem

[The Common Fund Data Ecosystem \(CFDE\)](#), a CDR program that is integral to the Common Fund's mission, aims to address challenges with data accessibility and reuse among Common Fund programs. In Phase 1, the program created an online discovery portal focusing on integrating Common Fund data resources into one location, making them findable, accessible, interoperable, and reusable (FAIR) for all users. The three primary goals for this phase include enabling users to utilize and query across multiple datasets, providing training and outreach, and integrating infrastructure and activities into a cohesive system. To achieve these goals, the ecosystem is structured so that individuals can act as both data consumers and data producers. Data Coordinating Centers (DCCs), an example of data producers, build the infrastructure necessary for data integration and reuse, ensuring that data remain accessible throughout a project and do not become obsolete over time.

Phase 2 runs from October 2023 to September 2028 and will focus on restructuring the program and developing five interconnecting centers: (1) Integration and Coordination Center, which coordinates between the other four centers and DCCs; (2) Knowledge Center, which focuses on integrating knowledge across Common Fund programs by establishing a knowledge portal; (3) Data Resource Center, which maintains the [CFDE portal](#), enabling users to query datasets; (4) Cloud Workspace Implementation Center, which will support cloud computing activities in select cloud workspaces beginning in fall 2024; and (5) Training Center, which will coordinate CFDE training efforts beginning in fall 2024.

CFDE focuses on building partnerships between DCCs to create opportunities for data to be standardized and available to all. Key resources in this effort include the Data Resource Center, which features the CFDE Workbench for searching across metadata, and the Knowledge Center, which offers a user context-driven Knowledge Portal. This portal provides access to a range of information across different scientific areas, helping users find relevant studies and resources from primary knowledge to tertiary levels. In addition, a cloud workspace will offer high-performance computing resources, increase access to data resources regardless of institutional size or location, and ensure support for both novice and expert users.

Training and outreach are also key priorities for the Common Fund and CFDE. Sub-initiatives include 2-year diversity fellowships to support trainees in using Common Fund data, with plans to award approximately five new training centers at universities. These centers will extend their reach to other institutions and communities to broaden participation. Starting in fall 2024, the Training Center will develop dynamic syllabuses and virtual mentoring, making educational materials globally accessible.

CFDE's success will be measured by its ability to facilitate new discoveries and enable a broad range of users to access and utilize Common Fund data. Upcoming plans include sharing information, offering training, and building events to drive uptake of Common Fund data, providing infrastructure and tools needed to enable broad data sharing, and creating a framework and resources to further plan for sustainability.

## **Panel Discussion: How Data Policy (Privacy and Ethics) Impacts Historically Underserved Communities**

*Moderator: Rachel Britt, Ph.D., OSC, NIH*

*Panelists:*

*Nyasha Chambwe, Ph.D., Feinstein Institutes for Medical Research*

*Mitchell Lunn, M.D., M.A.S., FACP, FASN, Stanford University*

*Krystal Tsosie, Ph.D., M.P.H., M.A., Arizona State University*

During this session, Dr. Britt posed several questions to the panelists, whose comments are summarized below.

### **Ensuring Representative Data in Research**

All of the panelists emphasized the importance of research datasets that are representative of the populations they aim to study. To ensure the inclusion of underserved and underrepresented communities in research, and to build trust with these communities, they should be engaged at every stage of the research process. Researchers must listen to input from community members when determining what data should be captured in a dataset. This practice is especially valuable for research in the lesbian, gay, transgender, queer, intersex, and asexual (LGBTQIA+) community, which is not adequately represented in the U.S. Census. Researchers must also consider the burden of data collection on study participants and the long-term usability of the data.

Dr. Tsosie added that the field of indigenous genomics faces unique challenges. Indigeneity is often conflated with race rather than recognized as a sociopolitical category. Researchers often prefer to draw reference group participants from a subset of tribes, rather than from the vast genetic, cultural, and sociopolitical diversity of the indigenous population. These reference groups may not serve as appropriate comparators for different indigenous communities in the United States. These researchers must reexamine whether their research questions align with the preferences and goals of the communities being studied because more context-specific questions may yield more relevant and informative data.

### **Accountability and Ethical Responsibilities in Research**

The panelists agreed on the need for a more comprehensive approach to data ethics and accountability in research, particularly concerning the rights and engagement of the communities involved. Dr. Tsosie highlighted the tendency of researchers to reuse previously collected genomic data, prioritizing their own research progress and reputation over continued community engagement, and neglecting the cultural norms and rights of the communities involved. To address this issue, academic programs should adopt a more holistic approach to training researchers in data ethics to equip them with the skills and responsibilities necessary to protect community rights. Dr. Chambwe agreed with this approach and emphasized the need for better data ethics training that extends accountability beyond funding agencies to include communities. Dr. Lunn added that accountability involves the use of both proactive measures (e.g., improved training, policy changes, technical assistance) and reactive measures, by which communities can assert their rights if ethical standards are violated, including public discourse or direct communication with research teams.

### **Transparency and Clarity in Data Use for Historically Underserved Communities**

The panelists acknowledged that historically underserved communities have developed a deep mistrust of health care and research because of historical injustices in research. Rebuilding their trust requires a transparent and intentional approach to data collection and use, as well as clear and meaningful communication about that approach. Researchers must explain how collected data will be used in ways that resonate with the communities involved to ensure that consent and understanding are truly achieved. Dr. Tsosie offered the Native BioData Consortium as an example of this approach; the consortium builds a data portal that addresses research questions of interest to the community and serves as an educational tool for community members. This resource builds community trust in research by enabling participants to access plain language summaries of studies using their data.

Dr. Lunn shared similar practical strategies for fostering transparency and trust within the LGBTQIA+ community. Dr. Lunn co-directs the [Population Research in Identity and Disparities for Equality \(PRIDE\) Study](#), which creates publicly available, plain language summaries of ongoing research studies to help participants understand how their data are being used. To protect participants' privacy, data viewers/users must submit an application that is reviewed by two scientific and community committees, each with equal input to ensure that data use aligns with community values.



### **Strategies for Ensuring Data Sovereignty and Data Justice**

Panelists highlighted the crucial role of robust data protections in ensuring that communities benefit from their participation in research. Dr. Tsoie called for careful consideration of data protections, particularly for non-federally recognized entities, and stressed the importance of adhering to ethical frameworks such as the FAIR Principles or the [Collective Benefit, Authority to Control, Responsibility, and Ethics \(CARE\) Principles for Indigenous Data Governance](#). Building on these principles, Dr. Lunn noted the importance of giving participants granular control over their data, allowing them to opt out or change how much of their data is publicly available. This capability is crucial for populations such as the LGBTQIA+ community, for which privacy is a primary concern.

Communities must be actively involved in both data collection and dissemination throughout the research process so that they can observe the benefits of their participation on an ongoing basis. As data methods continue to evolve, so should the tools that help them to engage with, understand, and interact with their data effectively. Promoting allyship is also essential, because the expertise of researchers from within minoritized communities is often undervalued or perceived as less objective. Researchers not from minoritized communities must, therefore, play a crucial role in amplifying and advocating for community voices.

### **Workshop Breakout Session: Community Engagement Strategies**

*Moderators:*

*Erin Beck, M.S., National Cancer Institute*

*Desiree Shantai Smith, Dr.PH., M.P.H.*

*Karriem Watson, DHSc, M.S., M.P.H., All of Us Research Program, NIH*

For this session, attendees were divided into small groups to discuss selected questions about relationship building and outreach efforts of the NIH Common Fund (see Appendix C for the full list of questions). Following the small group discussions, group members shared summaries of their group's conversations with all attendees present; findings from across the small breakout groups are summarized below.

#### **Best Practices for Community Engagement**

To establish community relationships, the Common Fund should engage with entities external to NIH. This effort might involve leveraging existing relationships and contacting community members or community-based organizations (CBOs), such as schools, religious communities, and tribal councils. When interacting with communities, establishing trust through transparency, authenticity, and consistency is paramount. By understanding the history of a community and listening to its members' experiences to understand their needs, researchers can build the foundation for genuine connections, identify true (not assumed) stakeholders, and ensure co-creation of research questions or projects. Providing funding should also be acknowledged as a best practice, even though systemic inequity affects how funding can be used (e.g., the indirect cost rate distribution between small CBOs and larger academic institutions).

To better sustain community relationships, the NIH Common Fund should directly engage with community members, empowering them to lead initiatives and offering opportunities for them to provide their feedback. Approaches toward this goal include hosting office hours and designating community advocates to serve as research team members with compensation for their time. Researchers must also ensure that their relationships with communities are cognizant of and resilient to change. For longer-term engagement beyond a period of funding or tenure, past awardees should be solicited for ideas about potential events, such as monthly webinars or ambassador referrals, that would extend their involvement with the NIH Common Fund. Hiring a permanent NIH Common Fund staff member whose role is focused on community engagement can help to advance these and other relationship building-oriented initiatives.

### **Leveraging Language for Engagement**

The NIH Common Fund can more effectively engage with communities by optimizing the use of language. Inclusive language in messaging to communities should be prioritized, and this messaging must be crafted to avoid reinforcing bias or stereotypes. In addition, researchers and communities may interpret research terms differently, and language *fluency* does not always directly translate to language *literacy*. Taking the time to create different versions of research materials that use community-accessible vocabularies can enhance communication. In addition, establishing a shared vocabulary between researchers and community members, potentially through a trusted messenger who understands the unique experiences of that community, can help to ensure that research language is better aligned with, and informed by, a community's identity.

### **Bidirectionality**

Attendees agreed that a bidirectional approach means that researchers identify commonalities and differences within a community's lived experiences that may require solutions unique to individual community members. Another way to prioritize community needs is by developing tools that focus on an effective user experience rather than the developers' needs, which may involve assessing community expertise about available resources and removing barriers to data access. Finally, relationships between research teams and communities should be established and maintained for the sake of relationship building, not for a pre-determined purpose or outcome. The NIH Common Fund should be self-aware and transparent about the motivations behind community engagement, with the goal of viewing these relationships as commitments rather than tasks to complete.

### **Outreach**

Attendees offered several suggestions to improve NIH Common Fund outreach efforts. Previous and current NIH Common Fund grantees could be surveyed to better understand their experiences using and accessing NIH Common Fund data resources. The NIH Common Fund could publicize information about upcoming events, new project results, and approaches to using available data through various communication channels to maximize audiences, including via Facebook because many communities regularly use this social media platform. Communications should use plain language to optimize community understanding of ongoing

research efforts that might be most impactful to them. One breakout group noted that short 10-minute videos tailored to a particular question or audience may be more effective than hosting large, generalized webinars. Finally, the NIH Common Fund should increase funding to directly support underrepresented populations and should consider how to engage Historically Black Colleges and Universities and other academic institutions without showing favoritism.

## **Panel Discussion: Data Use and Accessibility**

*Moderator: Maryam Zaringhalam, Ph.D., NLM, NIH*

*Panelists:*

*Seth Berke, Johns Hopkins Bloomberg School of Public Health*

*Frank Elavsky, Carnegie Mellon University*

*Katherine Kim, Ph.D., M.P.H., M.B.A., FAMIA, University of California, Davis School of Medicine and University of California, Berkeley*

During this session, Dr. Zaringhalam posed several questions to the panelists, whose comments are summarized below.

### **Data Value Continuum**

Dr. Kim noted that the value of data exists on a continuum that spans data to information to knowledge to wisdom. Data are variables, information places those variables in context through analyses, knowledge applies that information to implement helpful approaches, and wisdom reflects how such knowledge changes or benefits the general population. Gaps in the effective use of data throughout the continuum can be best addressed by answering questions regarding what data are collected and transmitted, how are they understood in the context of a certain project or initiative, and how they are effectively communicated to and digested by intended audience, and, more broadly, how this effort translates to meaningful change in the field at large.

### **Data Communication Strategies**

Panelists recognized the importance of effective communication when sharing and presenting data, which has implications for researchers as well as lay audiences. Dr. Kim noted that researchers have a responsibility to build competency and expertise in communicating not only the results of data analyses but also the contents of the data (e.g., how they are structured, analyzed, and managed) to non-research team members. Panelists also agreed that data presentations offer an opportunity to initiate and build relationships and that data visualization techniques can help researchers effectively communicate their data to various audiences. Mr. Berke suggested that communities receiving data should be provided with training and technical assistance focused on relevant project terminology to improve understanding of the data's purpose and how to access and navigate any related data sources.

### **Equitable Data Accessibility for People with Disabilities**

Mr. Berke suggested that curating data visualizations to incorporate colorblind-friendly color palettes and large, legible subtitles can make data more accessible to users with disabilities or other needs. In addition, panelists agreed that the community should be involved in

determining and addressing accessibility needs. Dr. Kim stated that accessibility is an equity issue and that the specific needs of each community should be considered when creating solutions and setting accessibility standards. Mr. Elavsky encouraged the audience to think about the meaning of the term “disability” and to frame research to support accessibility around not only people experiencing disabilities but also the public because people can become disabled by health, environmental, or other circumstances. Panelists also proposed involving community members in co-designing or co-producing accessible solutions that are adaptable to changing circumstances, and Mr. Elavsky added that these individuals should be compensated for their contributions. Mr. Elavsky also suggested consideration of the work already being done by these communities when creating new projects, especially to prevent re-invitation of labor from people with disabilities.

### **Prioritizing Section 508 Compliance**

[Section 508 of the Rehabilitation Act of 1973](#) is a federal law that outlines a set of standards for providing equal accessibility to information and services and requires U.S. government entities to reach a level of compliance with these standards. However, many organizations view these standards as optional rather than mandatory. To shift this perspective, Mr. Elavsky highlighted the importance of involving individuals with a thorough understanding of Section 508 standards in project working groups. Mr. Elavsky also noted that, although ground-up efforts dedicated to 508 compliance should be commended and supported, federal agencies ultimately drive commitment to 508 compliance.

### **Data Access Considerations**

Dr. Kim recognized the difficulty of procedurally adjudicating the extent to which datasets should be openly accessible. Many projects collect multimodal data because they provide comprehensive information about a community and its needs. However, such extensive data collection introduces the risk of reidentification, that is, the more data obtained about a specific group of people, especially in smaller, underserved communities such as rural, tribal, and island populations, the greater the risk of individuals in that group being identified and exposed to outside groups. Regulations such as the [Common Rule](#), which serves as the ethical standard in government funded human subjects research, do not always adequately address how to balance the risks with the opportunities, so decisions about data collection should be made in consultation with members of the subject community.

### **Resources for Data Use**

Mr. Berke highlighted the utility of cloud platforms for data use. Each platform offers a unique set of features and materials ranging from downloadable raw data files to highly structured tools that can be leveraged to analyze data in various ways. Panelists observed that end users select platforms based on their desire for tool and data analysis customization, computational experience, and accessibility needs. Effective use of cloud platforms requires an understanding of the data and tools being offered, and thus Mr. Berke suggested that each platform support users either through training on the data analysis tools or establishment of a coordinating center that processes and publishes data at various levels of granularity that can then be used for further analysis.

## Workshop Breakout Session: Data Challenges and Barriers to Data Use

*Moderators:*

*Lauren Amos, Ph.D., Mathematica*

*Samson Gebreab, Ph.D., M.Sc., ODSS, NIH*

*Avi Ma'ayan, Ph.D., Icahn School of Medicine at Mount Sinai*

For this session, attendees were divided into small groups to consider select questions about barriers to data access and use of data resources developed with NIH Common Fund support (see Appendix C for the full list of questions). Following the small group discussions, group members shared summaries of their group's conversations with all attendees present; findings from across the small breakout groups are summarized below.

### Barriers to Data Use

Although data use differs for each user and audience, attendees identified several common barriers to using data collected through NIH Common Fund initiatives. In general, accessing data can be difficult. Considerations about data safety and ethics related to equity and consent may prevent or limit release of certain datasets. Further, many communities involved in NIH Common Fund projects may lack reliable internet connections required to access data sources regardless of whether they are publicly available. In addition, many financial considerations affect data access; large datasets are expensive to store; some datasets are not accessible without grant funding; and low-resource institutions may be unable to secure competitive NIH funds to access these datasets.

Even when data sources are readily accessible, users may not know how to effectively utilize them. Many data systems are designed without input or feedback from potential end users, or when feedback mechanisms exist, comments are neither effectively collected nor responded to in a timely manner. Datasets that are accessible to the public are often tailored for researchers rather than the community members the data aim to support. Several other challenges can affect data use by end users and introduce the risk of uninformed decision-making. For example, cryptic data annotations not written in plain language prevent user identification of relevant information; difficulty navigating online interfaces might hide data of interest; and lack of understanding of data standards may obscure how these data can be used or interpreted.

Finally, information gaps exist in communities that could benefit from these data and related tools. Some communities distrust researchers, who often derive greater benefit from the data generated than the community members who are the focus of the study. In addition, communities often lack awareness of the concept and purpose of biomedical research as well as the resources that are already available to them.

### Strategies to Overcome Barriers to Data Use

Attendees provided several suggestions to overcome barriers associated with data access and use. First, platforms should prioritize the user experience, for example, by designing a user-friendly interface and offering instructional resources such as standardized protocol documents, tutorial videos, interactive graphics, informational webinars, and hands-on

workshops. Training could be tailored to engage different audiences, ranging from conventional researchers to researchers at low-resource institutions to nontraditional researchers such as community organizations, and could be led by mentors or individuals trained to use data resources of interest. Community-focused trainings could highlight the purpose and relevance of data resources, perhaps by incorporating a process that matches available data to audiences that can use them the most.

Feedback mechanisms should be developed to identify the most pertinent approaches to improving data access and use. Enabling end users to provide their feedback promotes inclusivity in the design of data sources and prioritizes topics and issues most relevant to communities of interest. The feedback mechanism should consider submission, compiling, tracking, and prioritization of, as well as timely response to, feedback provided from both internal and external NIH users. Creating a centralized service hub that assigns tickets to responses can help streamline this process.

Finally, outreach to directly connect NIH Common Fund researchers with communities can provide additional support for data access and use. Attendees recommended that the NIH Common Fund perform a landscape analysis to determine gaps in audiences engaging with available resources. Communities identified through this process, as well as low-resource institutions or small universities without research programs, can then be contacted and offered support, especially financial support, which could be drawn from existing mechanisms or new funding opportunities in the form of small grants. Attendees also noted that precision medicine and artificial intelligence can also be leveraged to close the gap in community awareness of and education on biomedical research and available related resources.

## Day 2: July 25, 2024

### Panel Discussion: Training & Education Needs of the Data Community

*Moderator: Nandita Rahman, All of Us Research Program, NIH*

*Panelists:*

*Jenea Adams, M.S., University of Pennsylvania*

*Kevin Cassel, Dr.PH., M.P.H., University of Hawaii Cancer Center*

*Zaki Sherif, Ph.D., Howard University College of Medicine and Georgetown University School of Medicine*

During this session, Dr. Rahman posed several questions to the panelists, whose comments are summarized below.

#### Best Practices for Data Literacy

Data literacy is multifaceted and may involve knowledge of the collection, management, statistical analysis, and visualization of data, as well as interpretation of analyses. All three panelists stressed that individuals whose data are collected should be empowered to understand their data and the potential influence of these data on policy and affecting change in their communities. Efforts should expand beyond making data available to community

members to making data applicable to their lives. Panelists also highlighted that training efforts should have practical and relevant outcomes for the community. Instead of using a generic problem set, trainers should use training datasets that help trainees to learn relevant or important aspects about their communities.

Many existing training programs in computational biology and data science focus on technically complex topics such as machine learning and artificial intelligence techniques. Despite the utility of these techniques, individuals could be trained in less complex topics, such as how to communicate about large-scale data. Ms. Adams also emphasized that training programs should equip trainees to meet the needs of their communities in addition to those of the research industry, a focus of her nonprofit organization, the [Black Women in Computational Biology Network](#).

Dr. Cassel emphasized the use of qualitative research methods to identify what data are important to communities and how they want to be trained. He added that training must happen on a personal level; trainees should know and have access to the trainer to ask questions and for assistance.

Data literacy audits within an organization provide a structured approach for self-assessing and seeking external feedback on data management, analysis, and interpretation skills. Researchers throughout the research pipeline, including faculty and career researchers, possess knowledge gaps; by systematically reviewing their skills via data literacy audits, these gaps can be identified and filled via tailored training programs. Some individuals, such as senior faculty, may be less willing to spontaneously acknowledge personal gaps. Thus, a standardized assessment or checklist that addresses gaps in all researchers may be most useful in improving data literacy.

### **Technology and Tools for Training and Education**

Dr. Sherif recommended the use of interactive dashboards, which allow users to gain experience with exploring and visualizing data. He noted that organizations, such as [The Carpentries](#), offer free workshops that focus on data visualization and dashboard creation. He noted that microlearning (i.e., the curation of material to be “bite-sized” for learners) and gamification (i.e., the practice of applying game style rewards, such as points, badges, and leaderboards to training materials) are promising approaches to promoting learner engagement, active participation, and retention.

Low-resource settings can pose challenges to provision of training. Dr. Cassel recommended the use of lower-tech solutions for these settings, including asynchronous learning through platforms such as Moodle and pre-recorded TED-style talks or lectures. He emphasized the need to teach community members how to access computer labs and other university-based resources, which may facilitate use of web-based training materials. Regardless of the method, Dr. Cassel stated that the communal aspect of learning, through use of trainee cohorts, is a concept to emphasize.

A wide variety of virtual training opportunities exists, and organizations may be able to provide funding to trainees to complete third-party trainings. However, because some individuals may

experience difficulty adhering to purely online asynchronous content such as Udemy or Coursera courses, Ms. Adams recommended the use of live courses in team settings when available. She stressed the importance of virtual spaces, such as Google Colab and Github, where trainees can practice using different software packages and datasets and where all-in-one controlled computing environments can be made easily available. She suggested that trainees should be ensured access to cloud computing tools, such as through cloud computing credits.

### **Making Outreach and Training Inclusive and Culturally Competent**

Ms. Adams described a program she codirected during her graduate training as an example of a successful training initiative. The program paid undergraduate students from community colleges and small colleges to attend a virtual internship experience in which they learned the fundamentals of computational biology and bioinformatics. The program paired student interns with mentors from similar backgrounds, and it focused on including students early in their degree programs so that they could alter their education plans based on the training. Most of the students who completed the program successfully moved into research positions, and some have started graduate school. Ms. Adams also shared that BWCB has an external mentorship program, named Connect Circles, focused on building community and sharing resources. Uniquely, this program aims to create bidirectional mentorship relationships, wherein both the mentor and mentee, regardless of background and expertise, contribute meaningfully to each other within the mentoring relationship.

Blended approaches, such as combining online modules with workshops or hackathons, might cater to diverse learning styles, leading to improved training outcomes. With these approaches, learners convene with trainers, ensuring quick responses to their questions and thus reducing frustration and increasing engagement and retention. Dr. Sherif promoted the use of community forums and mentorship programs within research communities to foster peer-to-peer learning opportunities and sharing of best practices.

### **Data Collection as an Area for Training**

Much of the effort to train communities focuses on data analysis and visualization; however, communities must also learn proper data collection methodologies. Many datasets are not hosted on shared repositories, such as those used by the NIH Common Fund, because the data were collected in a manner that complicates their use for secondary analysis. A lack of attention to data standards, the collection of metadata, and maintenance and publication of data dictionaries likely contribute to this issue; panelists agreed that these areas require additional training in the community.

Dr. Cassel provided an example of a community in American Samoa that had insufficient screening of and high mortality rates for colorectal cancer. His team provided training in responder-driven sampling data collection so that the community could track rates of colorectal cancer and assess health literacy in the community. As another example, he described a community whose participation in the [Health Information National Trends Survey \(HINTS\)](#) was hindered by low response rates to phone calls from project surveyors. Again, his team provided



training in responder-driven sampling, which improved data collection and enabled inclusion of the community's data in the national HINTS dataset.

### **Responsibilities of Data Analysis Tool Developers**

In response to a question, Dr. Sherif stated that data analysis tool developers should be responsible for defining guidelines for training and for integration of safety mechanisms into their systems to ensure appropriate data use, especially because data analysis and data cleaning may involve the use and modification of private patient data.

## **Workshop Breakout Session: Data Solutions to Identified Barriers from Day 1**

*Moderators:*

*Lauren Amos, Ph.D., Mathematica*

*Samson Gebreab, Ph.D., M.Sc., NIH*

*Avi Ma'ayan, Ph.D., Icahn School of Medicine at Mount Sinai*

For this session, attendees were divided into small groups to consider select questions about possible solutions to data access barriers identified during a breakout session on Day 1 of this workshop (see Appendix C for the full list of questions). Following the small group discussions, group members shared summaries of their group's conversations with all attendees present; findings from across the small breakout groups are summarized below.

### **Gathering Information on Communities and Their Needs**

Identifying spaces, both real and virtual, where community members can congregate will aid in defining communities, assessing their needs, and identifying individuals for educational initiatives. Community-based forums may enable researchers and educators to crowdsource research questions and obtain community buy-in.

Key stakeholders and champions within the community may provide feedback to help tailor educational trainings to meet community needs. Such informants may also aid in determining the immediate, intermediate, and long-term needs of the community, which will inform decisions about the material covered. For example, will a community need to ultimately apply for federal funding to solve the problems identified by the research project?

### **Maximizing Access to Resources in Community-Based Education**

When working with study participants and patient communities, training materials should be simple, authentic, engaging, and introductory. Conversations between researchers and the community should discuss participant involvement in the research process, data ownership, and the participants' feeling about the data collected and how those data will be used. Researchers may also consider participant-facing portals where community members can see how their data are being used and gain access to study data following study completion. All such participant-facing digital resources should be accessible on mobile devices and compliant with relevant accessibility standards.

Once an initial educational initiative is complete, mentors and educators should return to the community to offer support, guidance, consultation, and mentorship. Train-the-trainer models may help to form peer support for community members.

### **The Ideal Role of a Community Partner**

Community advisory board members should have a larger role in directing research across the entire data pipeline. Community members know the important barriers relevant to the community and outcomes that would be desired from research initiatives. Advisory boards should be constructed to promote engagement and the delivery of meaningful guidance. Different roles could be developed for community partners, allowing for different levels of involvement and in different domains—including program-level partners, educational partners, and networking partners.

An important step to working within a community is determining the relevant and necessary community partners, which could include local health agencies, health care systems, religious leaders, and public-school representatives.

### **Shifting Toward Meaningful Co-Development of Research with Community Partners**

To meaningfully co-develop research with researchers, community partners should be involved from the beginning of the funding process, including as grant reviewers and IRB members. This approach helps to ensure that funded and approved research addresses the actual needs of the community and maximizes the potential of the research to benefit the community. When an appropriate contribution has been made, community members should be included as authors on scientific manuscripts.

Researchers may consider a strategic partnership plan or community engagement plan, wherein objectives of the partnership are tracked throughout the research process, and research findings and relevant knowledge are transferred from the researcher to the community. Such plans should be standardized.

Involvement of community members in research consortia may formalize the role of the community and link community involvement to funding. Efforts should also be made to build enough capacity within communities that they can independently apply for funding and drive research initiatives that directly impact them. Achieving this level of community involvement in the research process would likely require systematic changes to current funding structures to enable these communities to be competitive for funding.

## Workshop Breakout Session: Immediate Training Needs and Supports of the Community

*Moderators:*

*Rachel Britt, Ph.D., OSC, NIH*

*Mackenzie Brandes, Broad Institute of Massachusetts Institute of Technology (MIT) and Harvard*

*Noël Burt, Broad Institute of MIT and Harvard*

For this session, attendees were divided into small groups to consider select questions about training needs and ways to support the community of researchers (including trainees) that use Common Fund data (see Appendix C for the full list of questions). Following the small group discussions, group members shared summaries of their group's conversations with all attendees present; findings from across the small breakout groups are summarized below.

### **Making NIH Common Fund Data More User-Friendly**

Attendees recommended that the NIH Common Fund create a more streamlined “digital front door” for its data and training resources, which should include an easier way to learn what data types are available and harmonized. If possible, the Common Fund's data-sharing platforms must have an intuitive and rational search system so that non-experts can identify which scientific questions can be answered using NIH Common Fund data. These intuitive searches should also allow users to identify relevant trainings and training materials. Instructions for accessing NIH Common Fund data resources should be easily understood by lay audiences and avoid jargon. Community members should be involved in vocabulary definition and materials generation efforts.

In its efforts to promote more widespread use of its data, the NIH Common Fund may consider more targeted outreach by identifying relevant communities; assessing their objectives for NIH Common Fund data use, the resources and technologies available to them, their mix of trainees versus career researchers; and then tailoring outreach to meet their needs. Individuals within these communities who have already used the resources can be engaged to act as within-community trainers and can be incentivized to do so via badges or certificates. Further, users should rate the quality of NIH Common Fund datasets based on, for example, the quality of included metadata, the data structure, and ease of use of the data.

Common Fund data used in publications are subject to Common Fund data standards, which may require significant work by data submitters to meet. Consequently, data are often submitted that do not meet Common fund standards, leading to data that are more difficult to use or access. Additional funding and support should be made available to incentivize the meeting of these standards.

### **Ideas for Providing Training and Educational Support to the Community**

In trainings in a community setting, community members who have experience with the platform or dataset can play an invaluable role. These individuals could be trained to help provide content expertise to fellow community members. They could record informational

videos, in which they share relevant information and pitfalls to avoid and lend perspective to new users. These individuals could receive credit for their efforts, perhaps through badges or certificates, which is a low-cost approach. In general, training organizations should perform an assessment within target communities to determine the resources already present so that they can be integrated into training plans.

Important considerations are where text or video content is made available and how that content is marketed. Spaces in which a community tends to access for training or educational content should be prioritized. If no such space exists, then a platform for sharing content should be developed. By implementing multiple strategies to deliver educational content, diverse learning styles within the same community can be accommodated. Some learners may prefer longer style presentations, such as full lectures or TED-style talks, whereas others with limited time may prefer concise but dense 10-minute clips. Video materials should be created with these diverse preferences in mind and should be edited to create different end products for different learners.

## **Workshop Closing Presentation & Wrap Up**

*Christopher Kinsinger, Ph.D., OSC, NIH*

Dr. Kinsinger discussed the polarity in the core mission of the NIH: “turning discovery into health.” Discovery is associated with the scientific investigation of universal truths, but health is not a universal idea: it often varies by community. Universal truths discovered scientifically must be contextualized in ways that will promote health in specific communities.

Dr. Kinsinger then highlighted several principles that emerged during discussions from this workshop:

- Use data for good. The meaning of “good” may differ based on community context, but data should be used for purposes that communities define as good.
- Transparency builds trust. Letting community members know what data are being collected and how they are being used is an important part of forming a trusting relationship with communities.
- The majority of the data life cycle lies in the storage and reuse phases. Thus, community engagement must continue into these phases.
- Community engagement should occur at every step of a research project.
- Collective benefit, authority to control, responsibility, and ethics constitute the CARE principles, which can be embedded in all research practices.

Dr. Kinsinger highlighted practices that OSC and this audience should consider as next steps:

- Data resources, communications practices, and training materials should be contextualized. Tailoring is important, but in the absence of real collaboration with the subject community, these practices alone are not enough.

- Data dissemination is bidirectional. OSC should expand its focus from what researchers want to investigate to what communities want to assess and better understand.
- In addition to initiatives that aim to provide data access to researchers, OSC should also consider other communities' data access needs and their expectations for these resources that are generated with their tax dollars.

Dr. Kinsinger commented that the presence of NIH staff members at this workshop gives him hope for the future of community data engagement. These individuals will be the ones writing funding opportunities and policies and having conversations with communities, thereby enhancing the potential for truly improved engagement. He added that this workshop represents the beginning of the NIH Common Fund's process toward community engagement goals.

Dr. Kinsinger asked meeting attendees to raise points for future consideration or key takeaways that they gained from the workshop:

- The role of health equity and health disparities in relation to data literacy and community engagement should be explored.
- Community engagement and partnership should be a part of research methodology prior to the formulation of research hypotheses and should be carried throughout the whole research process. Approaches for successfully doing this should be standardized and systematized.
- The administrative and logistical hurdles present in the current grant systems inhibit the effective involvement of community members in research.
- The new Office of Management and Budget revisions to Statistical Policy Directive Number 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity may be relevant to these discussions.
- In relation to education and training, more research groups should co-design training with the communities with which they collaborate.
- Many communities are technically sophisticated in their own right; thus, community educational initiatives should not always assume a total lack of data literacy.
- The gamification of training materials, such as the use of badges or a leaderboard, seems particularly promising.

## Appendix A: Agenda

### Day 1: July 24, 2024

- 9:00**            **Welcome and Introduction**  
Douglas Sheeley, Sc.D., OSC, NIH  
Natalie Vineyard, M.S., OSC, NIH
- 9:15**            **Keynote Presentation: Revisiting Preconceived Notions of Community Engagement**  
Krystal Tsosie, Ph.D., M.P.H., M.A., Arizona State University
- 10:00**           **NIH Common Fund and Common Fund Data Ecosystem Resources**  
George Papanicolaou, Ph.D., OSC, NIH
- 10:30**           **Break**
- 10:45**           **Panel Discussion: How Data Policy (Privacy and Ethics) Impacts Historically Underserved Communities**  
Moderator: Rachel Britt, Ph.D., OSC, NIH  
Panelists: Nyasha Chambwe, Ph.D., Feinstein Institutes for Medical Research  
Mitchell Lunn, M.D., M.A.S., FACP, FASN, Stanford University  
Krystal Tsosie, Ph.D., M.P.H., M.A., Arizona State University
- 12:00**           **Lunch**
- 1:00**            **Workshop Breakout Session: Community Engagement Strategies**  
Moderators: Erin Beck, M.S., NCI  
Desiree Shantai Smith, M.P.H., DrPH  
Karriem Watson, DHSc, M.S., M.P.H., *All of Us* Research Program, NIH
- 2:00**            **Panel Discussion: Data Use and Accessibility**  
Moderator: Maryam Zaringhalam, Ph.D., NIH  
Panelists:  
Seth Berke, Johns Hopkins Bloomberg School of Public Health  
Frank Elavsky, Carnegie Mellon University  
Katherine Kim, Ph.D., M.P.H., M.B.A., FAMIA, University of California, Davis School of Medicine and University of California, Berkeley
- 3:00**            **Break**
- 3:15**            **Workshop Breakout Session: Data Challenges and Barriers to Data Use**  
Moderators:  
Lauren Amos, Ph.D., Mathematica

Samson Gebreab, Ph.D., M.Sc., NIH  
Avi Ma'ayan, Ph.D., Icahn School of Medicine at Mount Sinai

**4:15 Closing Remarks**

**4:30 Adjourn Day 1**

## **Day 2: July 25, 2024**

**9:00 Panel Discussion: Training & Education Needs of the Data Community**  
Moderator: Nandita Rahman, *All of Us* Research Program, NIH  
Panelists: Jenea Adams, M.S., University of Pennsylvania  
Keven Cassell, Dr.PH., M.P.H., University of Hawaii Cancer Center  
Zaki Sherif, Ph.D., Howard University College of Medicine and Georgetown University School of Medicine

**10:00 Concurrent Workshop Breakout Sessions**

### **Data Solutions to Identified Barriers from Day 1**

Moderators: Lauren Amos, Ph.D., Mathematica  
Samson Gebreab, Ph.D., M.Sc., NIH  
Avi Ma'ayan, Ph.D., Icahn School of Medicine at Mount Sinai

### **Immediate Training Needs and Supports of the Community**

Moderators: Rachel Britt, Ph.D., OSC, NIH  
MacKenzie Brandes, M.B.A., Broad Institute of MIT and Harvard  
Noël Burt, Broad Institute of MIT and Harvard

**11:20 Break**

**11:30 Workshop Closing Presentation & Wrap Up**  
Christopher Kinsinger, Ph.D., OSC, NIH

**12:30 Adjourn Day 2**

## Appendix B: Meeting Speakers and Personnel

### Speakers, Panelists, and Moderators

Jenea Adams, University of Pennsylvania  
Lauren Amos, Mathematica  
Erin Beck, NCI, NIH  
Seth Berke, Johns Hopkins Bloomberg School of Public Health  
MacKenzie Brandes, Broad Institute of MIT and Harvard  
Rachel Britt, OSC, NIH  
Noël Burt, Broad Institute of MIT and Harvard  
Kevin Cassell, University of Hawaii Cancer Center  
Nyasha Chambwe, Feinstein Institutes for Medical Research  
Frank Elavsky, Carnegie Mellon University  
Samson Gebreab, ODSS, NIH  
Katherine Kim, University of California, Davis School of Medicine and University of California Berkeley  
Christopher Kinsinger, OSC, NIH  
Mitchell Lunn, Stanford University  
Avi Ma'ayan, Icahn School of Medicine at Mount Sinai  
Toy Lisa Mitchell, OSC, NIH  
George Papanicolaou, OSC, NIH  
Nandita Rahman, *All of Us* Research Program, NIH  
Douglas Sheeley, OSC, NIH  
Zaki Sherif, Howard University College of Medicine and Georgetown University School of Medicine  
Desiree Shantai Smith  
Krystal Tsosie, Arizona State University  
Natalie Vineyard, OSC, NIH  
Karriem Watson, *All of Us* Research Program, NIH  
Maryam Zaringhalam, NLM, NIH

### Meeting Organizers

Andréa Harris, OSC, NIH  
Cheryl Mavritte, OSC, NIH  
Christopher Kinsinger, OSC, NIH  
Karen Kellton, OSC, NIH  
Natalie Vineyard, OSC, NIH  
Rebecca Black, OSC, NIH  
Rachel Britt, OSC, NIH  
Richard Conroy, OSC, NIH



Sahana Kukke, OSC, NIH  
Toy Lisa Mitchell, OSC, NIH

**Special Thanks**

Andy Burnim, OSC, NIH  
Kaitlyn Browning, OSC, NIH  
Katelynn Milora, OSC, NIH  
Kristen Schlotman, OSC, NIH  
Michelle Schneider, OSC, NIH  
Vanessa Barnes, OSC, NIH

## Appendix C: Workshop Breakout Sessions Discussion Questions

### Community Engagement Strategies

#### *Relationship Building*

- What are the best practices and strategies of initiating, building, and sustaining relationships with communities?
- What are the best ways to center the use of language in relationship building and community engagement?
- What are the characteristics of bidirectionality with community engagement?
- What are the mistakes or pitfalls that have occurred in the past that should be avoided when we seek to build relationships with the community?
- What are the best practices that we should use when cultivating relationships with the community?

#### *Outreach*

- In what ways can the Common Fund demonstrate more inclusivity in its outreach efforts?
- What are the most effective forms of communication and outreach as we engage with the community?
- Outreach to underrepresented and smaller universities, Minority Serving Institutions, and Historically Black Colleges and Universities
- Uses of social media and other forms of communication: what and who are we missing?
- What are the mistakes or pitfalls that have occurred in the past that should be avoided when we engage in outreach with the community?
- What are the best practices that we should use when reaching out to the community?

### Data Challenges and Barriers to Data Use

- What are the barriers to using NIH-supported data resources like those developed by Common Fund programs?
- What are some strategies the Common Fund should consider (such as funding, training, and educational strategies, etc.) to address challenges/barriers of data use?

### Immediate Training Needs and Supports of the Community

- What ways can the NIH Common Fund make data more approachable to the research community and trainees?
- How can NIH Common Fund foster understanding of datasets and their potential?
- Data scientists and computer scientists like working with data, but clarify what else that can be done that would be helpful? Is there something to focus on or specific barriers that can be removed to support the community?

- What is the best way to provide training and educational support to the community (e.g. free informational webinars, on demand trainings, in-person trainings, office hours, peer learning workshops, mentorship programs, toolkits, etc.)?

## Appendix D: Website Links

All of Us	<a href="https://allofus.nih.gov/">https://allofus.nih.gov/</a>
Black Women in Computational Biology Network	<a href="https://www.blackwomencompbio.org/">https://www.blackwomencompbio.org/</a>
Common Rule	<a href="https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html">https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html</a>
Collective Benefit, Authority to Control, Responsibility, and Ethics (CARE) Principles for Indigenous Data Governance	<a href="https://www.gida-global.org/care">https://www.gida-global.org/care</a>
Health Information National Trends Survey (HINTS)	<a href="https://hints.cancer.gov/">https://hints.cancer.gov/</a>
Indigenous Data Science Education (IndigiData)	<a href="https://indigidata.org/home/">https://indigidata.org/home/</a>
Native BioData Consortium (Native BioData)	<a href="https://nativebio.org/">https://nativebio.org/</a>
Population Research in Identify and Disparities for Equity (PRIDE) Study	<a href="https://pridestudy.org/">https://pridestudy.org/</a>
Rapid Acceleration of Diagnostics (RADx)	<a href="https://www.nih.gov/research-training/medical-research-initiatives/radx">https://www.nih.gov/research-training/medical-research-initiatives/radx</a>
Section 508	<a href="https://www.section508.gov/manage/laws-and-policies/">https://www.section508.gov/manage/laws-and-policies/</a>
The Carpentries	<a href="https://carpentries.org/">https://carpentries.org/</a>
The Common Fund Data Ecosystem (CFDE) Portal	<a href="https://info.cfde.cloud/">https://info.cfde.cloud/</a>
The NIH Common Fund	<a href="https://commonfund.nih.gov/">https://commonfund.nih.gov/</a>
The NIH Common Fund Data Ecosystem (CFDE)	<a href="https://commonfund.nih.gov/dataecosystem">https://commonfund.nih.gov/dataecosystem</a>