NATIONAL INSTITUTES OF HEALTH

NIH EFFORTS IN SUPPORT OF OPEN DATA

An Integrated Approach

Lyric Jorgenson, PhD
Acting NIH Associate Director for Science Policy
Acting Director of the Office of Science Policy

Patricia Brennan, RN, PhD, FAAN
Director, National Library of Medicine

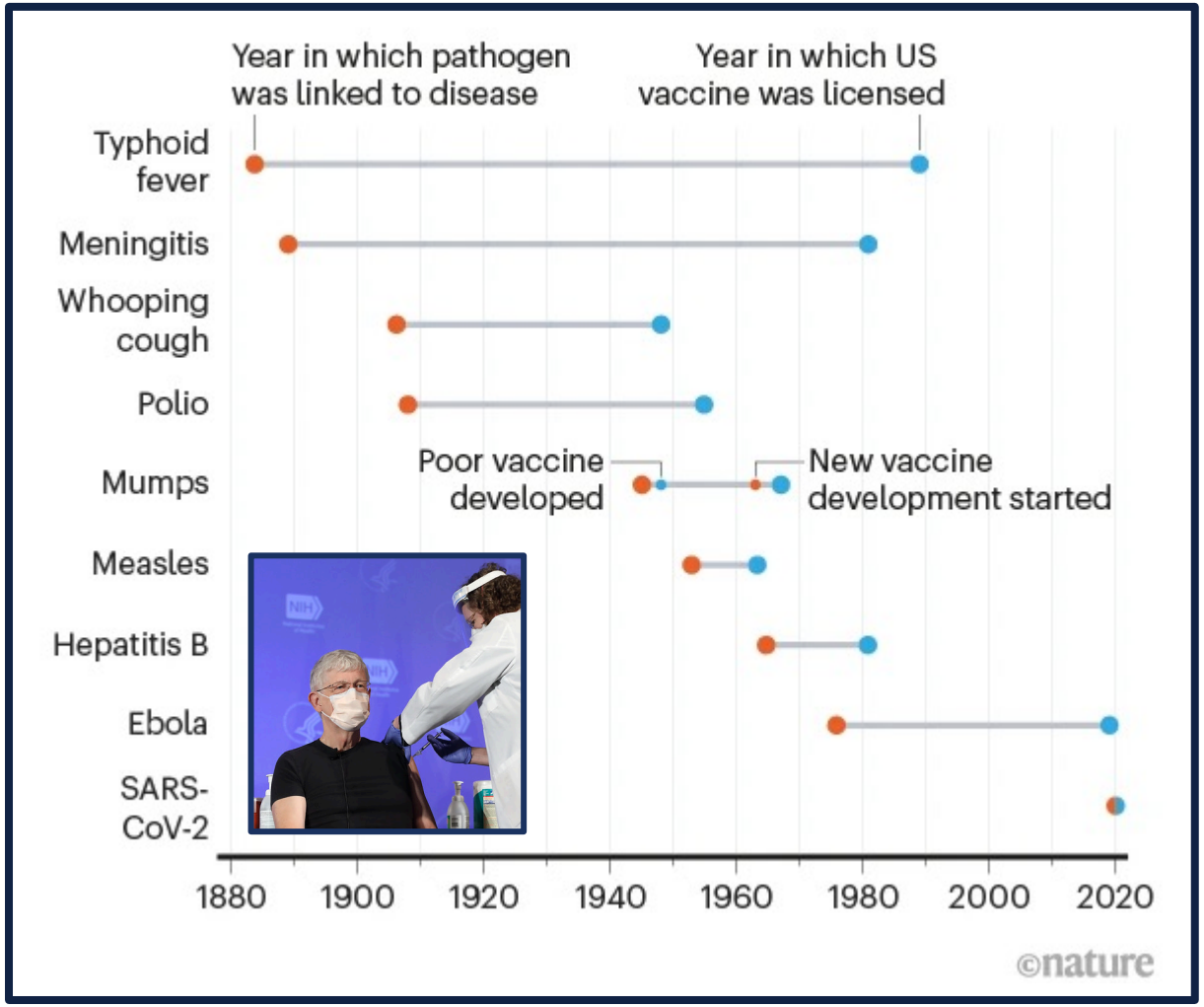
Susan Gregurick, PhD
Associate Director for Data Science
Director of the NIH Office of Data Science Strategy

Mike Lauer, MD
Deputy Director for Extramural Research

ACD Meeting December 9th, 2022

MEETING HEALTH NEEDS THROUGH DATA SHARING

OPEN DATA TO ADDRESS THE COVID-19 PANDEMIC



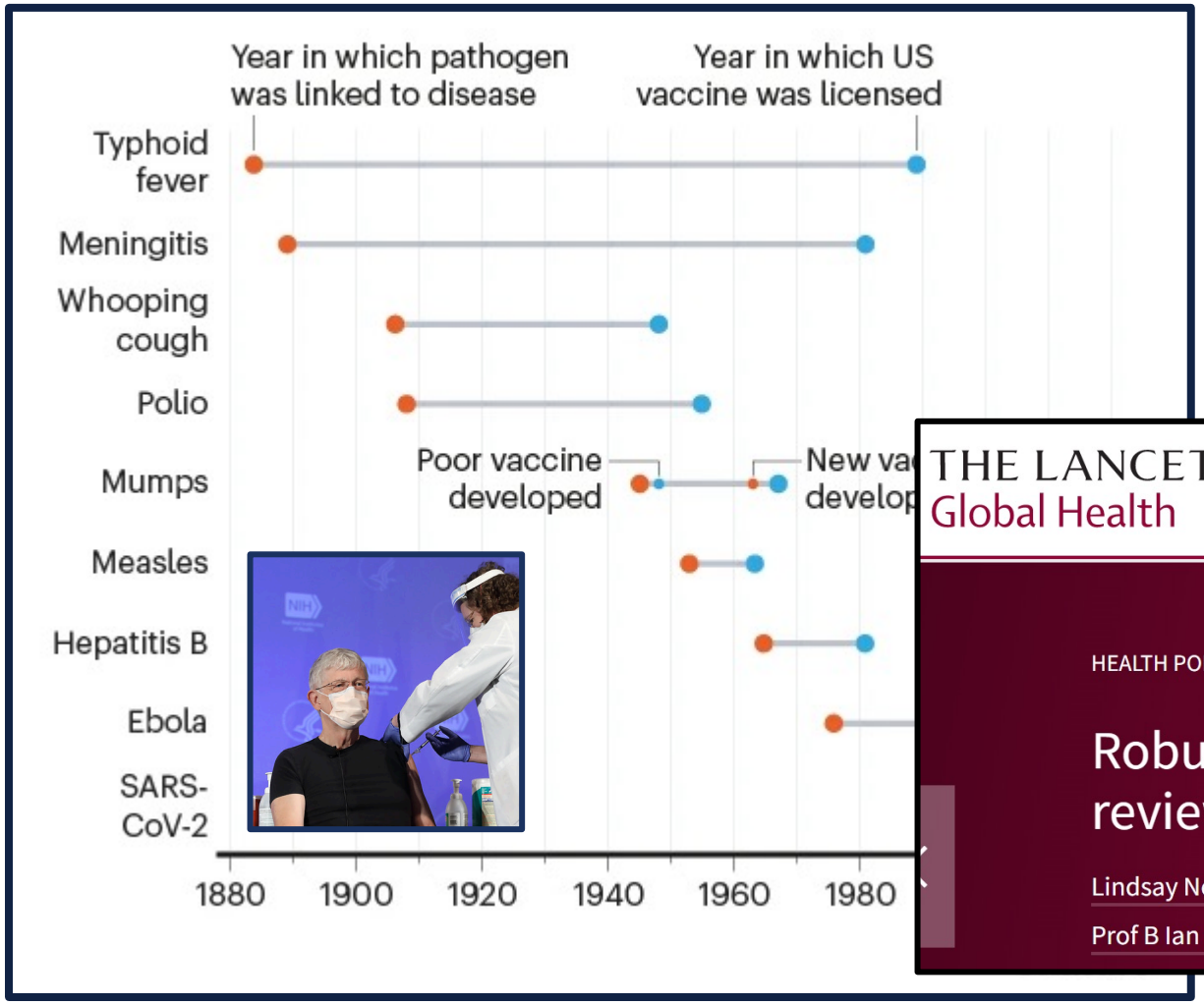
Pandemic sharing and immediate access to:

- > 2.5M SARS COV-2 genomic sequences
- > 1,300 SARS-CoV-2 protein structures
- > 300 reagents for biomedical research
- > 7.3B rows of clinical data (N3C)
- > 150K papers

Source: niaid.nih.gov

MEETING HEALTH NEEDS THROUGH DATA SHARING

OPEN DATA TO ADDRESS THE COVID-19 PANDEMIC



Pandemic sharing and immediate access to:

- > 2.5M SARS COV-2 genomic sequences
- > 1,300 SARS-CoV-2 protein structures
- > 200 reagents for biomedical research

A GOAL WE CAN AGREE ON

MAXIMIZING THE VALUE OF DATA

Scientific data are the catalyst for breakthroughs and discoveries

Integrated policies, resources, and infrastructure are key to responsible sharing and reuse

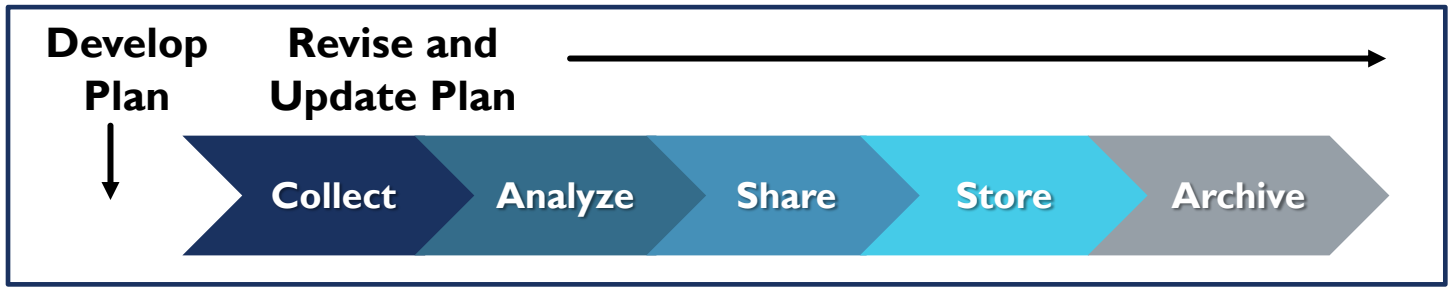
NIH aims to promote effective data sharing as the rule, not the exception

NIH VISION

AN INTEGRATED AND OPEN DATA ECOSYSTEM

NIH DATA MANAGEMENT & SHARING POLICY

Prospective planning and management will…



- ✓ **Spur faster, higher quality scientific results**
- ✓ **Encourage greater scientific integrity**
- ✓ **Enable future inquiry, discovery, and translation**

MANAGEMENT AT ALL STAGES

A VIRTUOUS CYCLE TO MAXIMIZE DATA UTILITY AND REUSE

FAIR Principles **Metadata** **Consent & Privacy** **Storage Costs** **Interoperability**

Data Standards **Curation Costs** **Security & Integrity** **Data Agreements**

Common Data Elements **Access** **Persistent Identifiers (PIDs)**

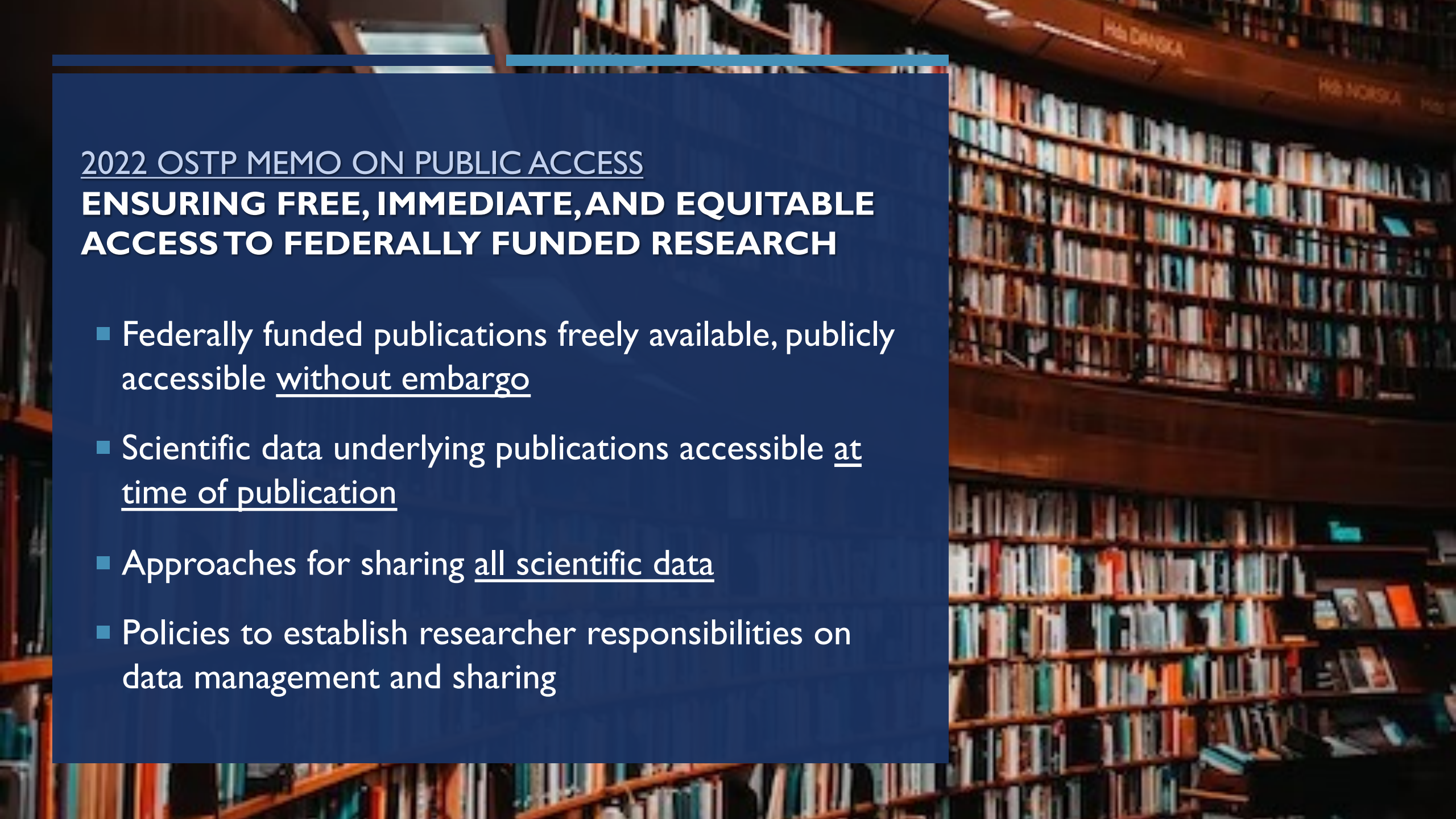
MANAGEMENT AT ALL STAGES

A VIRTUOUS CYCLE TO MAXIMIZE DATA UTILITY AND REUSE

FAIR Principles **Metadata** **Consent & Privacy** **Storage Costs** **Interoperability**

Data Standards **Curation Costs** **Security & Integrity** **Data Agreements**

Common Data Elements **Access** **Persistent Identifiers (PIDs)**



2022 OSTP MEMO ON PUBLIC ACCESS

ENSURING FREE, IMMEDIATE, AND EQUITABLE ACCESS TO FEDERALLY FUNDED RESEARCH

- Federally funded publications freely available, publicly accessible without embargo
- Scientific data underlying publications accessible at time of publication
- Approaches for sharing all scientific data
- Policies to establish researcher responsibilities on data management and sharing

NIH STATEMENT

SUPPORT FOR PUBLIC ACCESS

August 25, 2022

Statement on NIH plans to speed access to federally funded research results

Today, the White House Office of Science and Technology Policy (OSTP) issued updated policy guidance directing federal agencies to expedite access to results of federally funded research. NIH has long championed principles of transparency and accessibility in NIH-funded research and supports this important step by the Biden Administration.

Over the coming months, NIH will work expeditiously to develop and share its plans for implementing the OSTP policy guidance. NIH intends to work with interagency partners and stakeholders to revise its current Public Access Policy to enable researchers, clinicians, students, and the public to access NIH research results immediately upon publication. I am pleased to report that NIH's efforts to maximize access to scientific data are already underway through implementation of the new NIH Policy for Data Management and Sharing (DMS Policy), which takes effect on January 25, 2023. Through the DMS Policy, NIH clearly articulates that sharing scientific data is fundamental to accelerating biomedical research discovery. Our DMS Policy implementation efforts continue, and I encourage you to visit sharing.nih.gov for the latest updates and resources that NIH has developed to support our community of researchers and institutions.

We are enthusiastic to move forward on these important efforts to make research results more accessible and look forward to working together to strengthen our shared responsibility in making federally funded research results accessible to the public.

Resources:

- OSTP Updates to Policy Guidance on Increasing Equitable Access to Federally Funded Research Results
- NIH Public Access Policy
- NIH Policy for Data Management and Sharing
- Sharing.nih.gov

Lawrence A. Tabak, D.D.S., Ph.D.
Performing the Duties of the NIH Director

NIH has long championed principles of transparency and accessibility in NIH-funded research and supports this important step by the Biden Administration

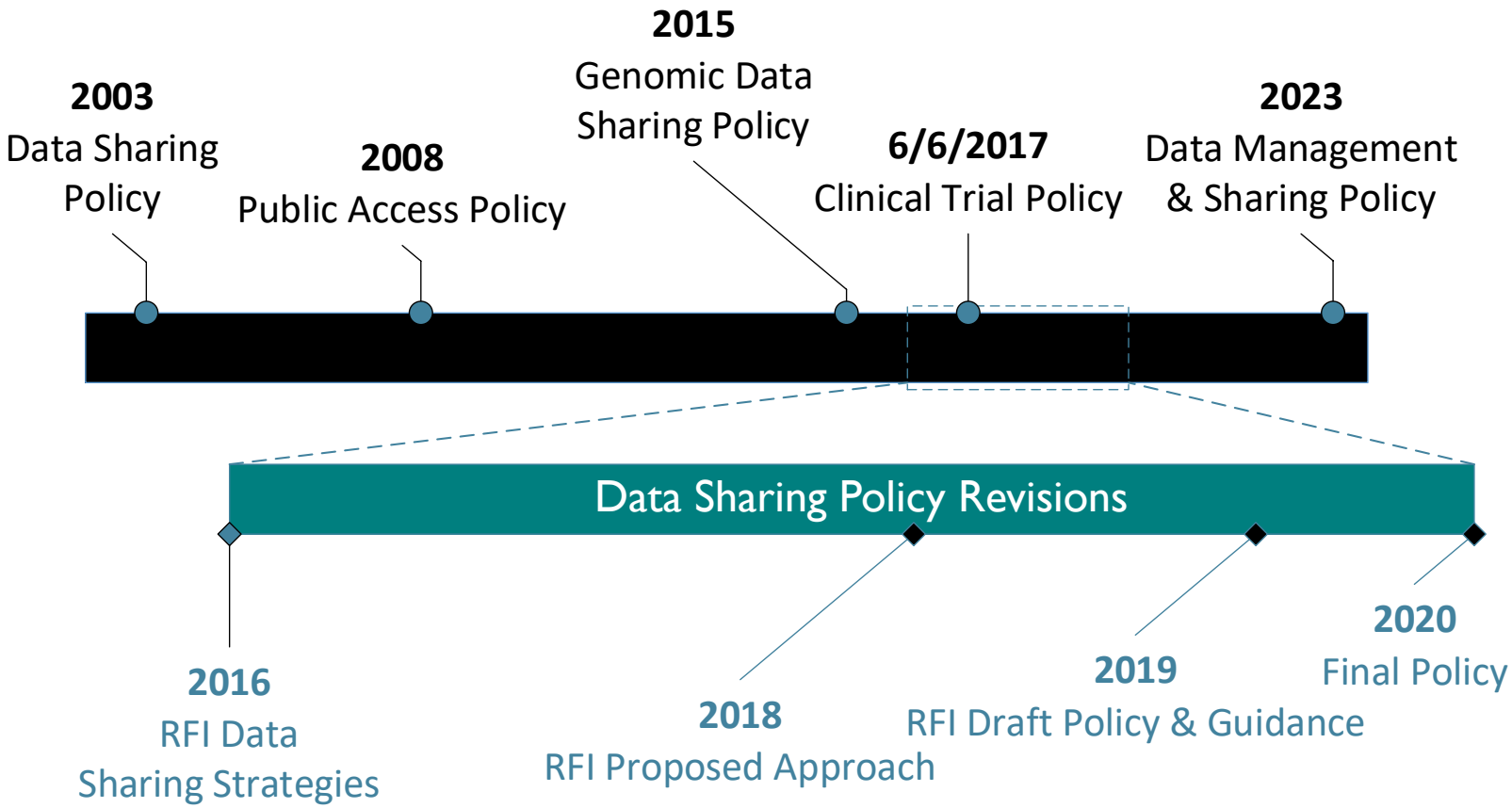
A SECOND GOAL WE CAN AGREE ON:

MINIMIZING THE BURDEN TO THE COMMUNITY

- We're building from a strong foundation
- Prospective planning will facilitate all stages of sharing and access
- We have developed resources, best practices, and guidance for the community

BUILDING FROM A STRONG FOUNDATION

DECADES OF NIH PUBLIC ACCESS AND DATA SHARING



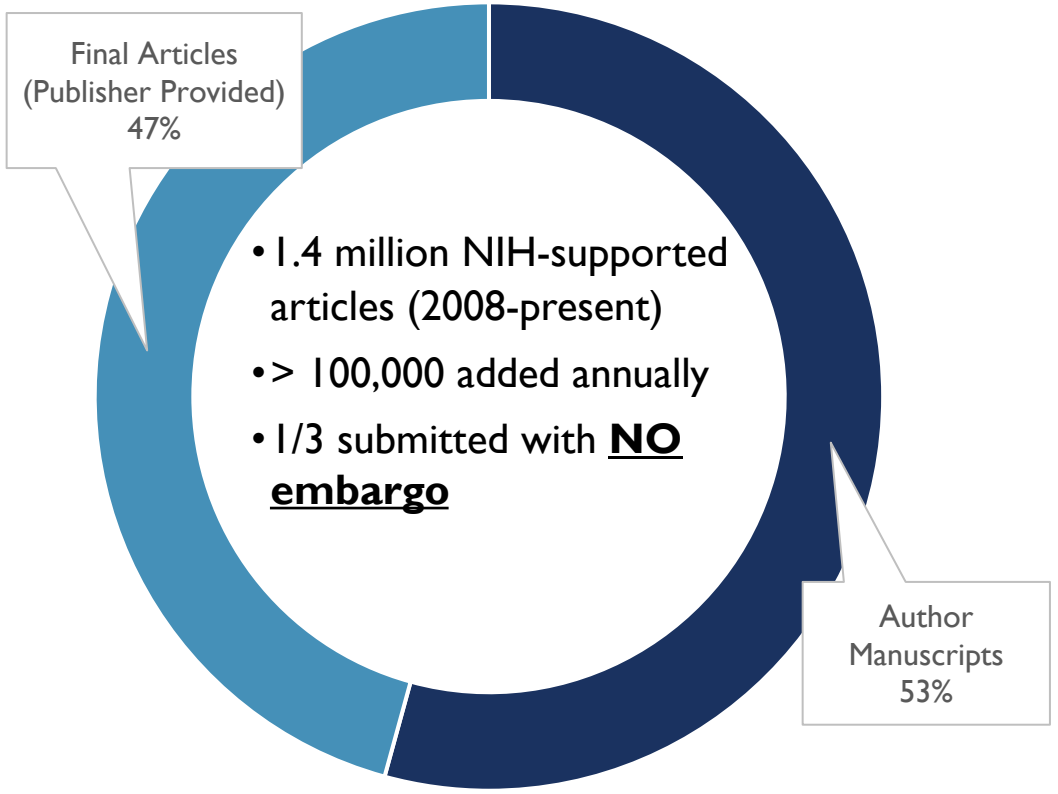
HOW IT STARTED, HOW IT'S GOING

SUCCESS OF CURRENT NIH PUBLIC ACCESS POLICY (AFTER 2008)

The Director of the National Institutes of Health shall require that ***all investigators funded by the NIH*** *submit or have submitted for them to…* ***PubMed Central*** *an electronic version of their final peer-reviewed manuscripts* ***upon acceptance for publication****, to be made* ***publicly available no later than 12 months after the official date of publication****…*

Implements Division G, Title II, Section 218 of PL 110-161
Consolidated Appropriations Act, 2008

NIH-Supported Articles in PMC

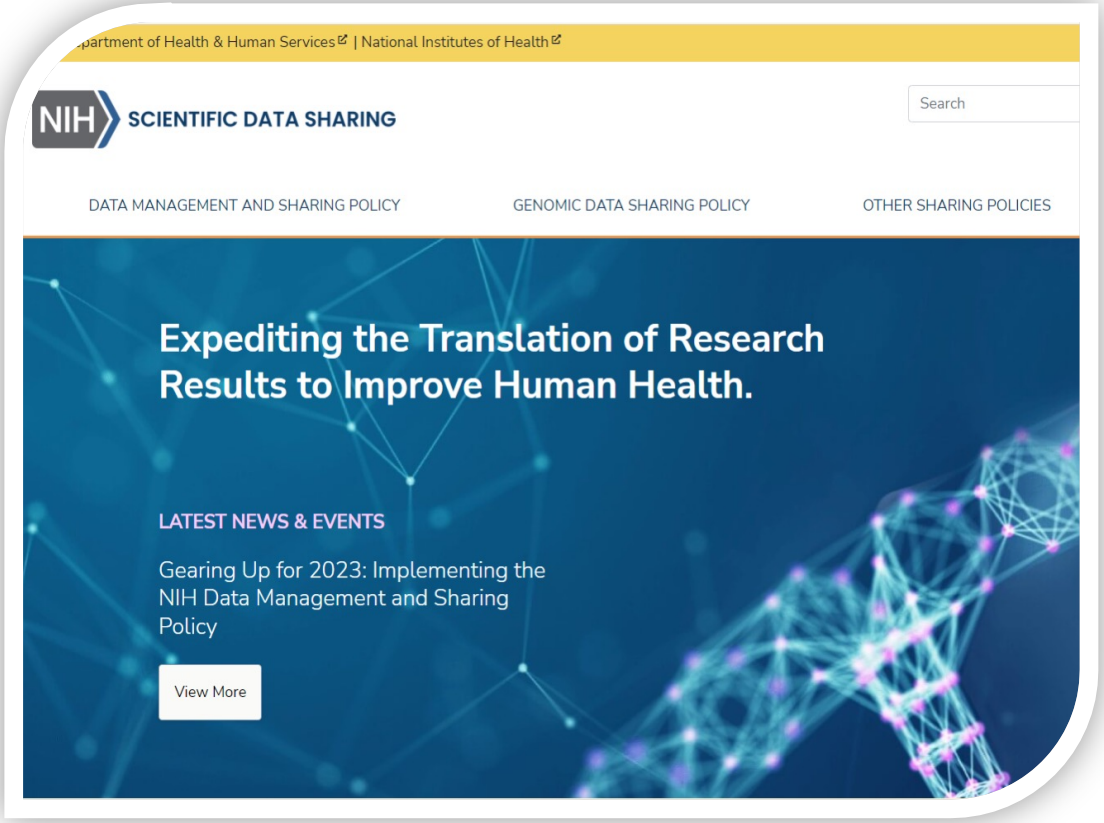


NIH POLICY FOR DATA MANAGEMENT & SHARING

- TWO BASIC REQUIREMENTS
  - Submission of a Data Management & Sharing "Plan" for all NIH-funded research
  - Compliance with the ICO-approved Plan
- Effective January 25, 2023 *(replaces 2003 Data Sharing Policy)*

RESOURCES TO HELP THE COMMUNITY

DATA MANAGEMENT AND SHARING



Resources available:

- Frequently Asked Questions
- Supplemental Information Documents:
  - Elements of a DMS Plan
  - Selecting a data repository
  - Allowable Costs
  - Protecting Privacy
  - Responsible Management/Sharing of American Indian/Alaska Native Participant Data
- Publicly posted webinars

Accessible at: sharing.nih.gov

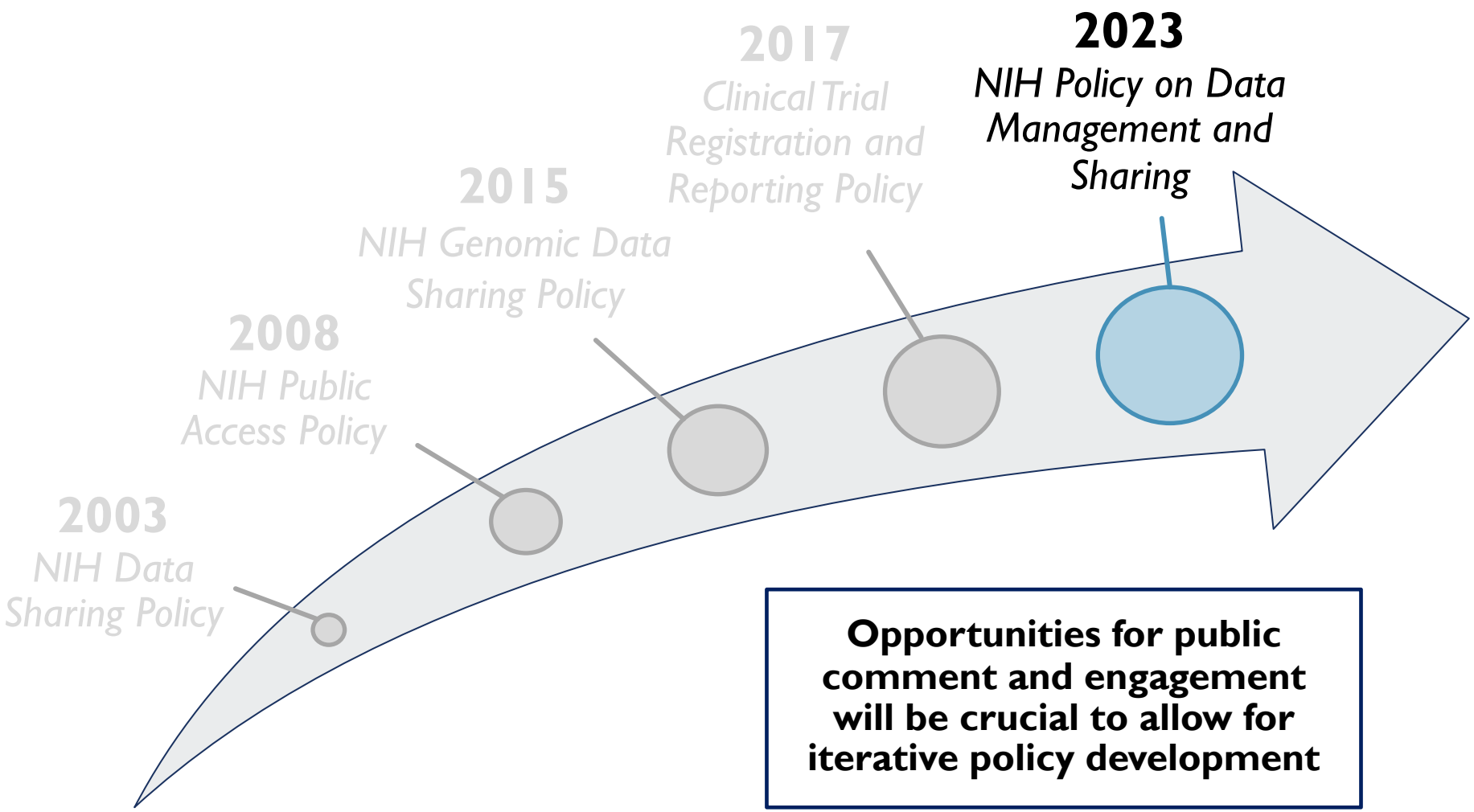
POISED FOR SUCCESS

NIH HAS MECHANISMS AND INFRASTRUCTURE IN PLACE

2022 OSTP Memo	Current approach, resources
▪ Federally funded publications freely available, publicly accessible without embargo	Public Access Policy (2008; to be revised) PMC PubMed Central®
▪ Scientific data underlying publications accessible at time of publication ▪ Approaches for sharing all scientific data ▪ Policies to establish researcher responsibilities on data management and sharing of research results (including metadata and persistent identifiers	Policy on Data Management and Sharing (2023; meets expectations of OSTP Memo) NIH SCIENTIFIC DATA SHARING sharing.nih.gov

BUILDING FROM A STRONG FOUNDATION

PLAN TO MEET PUBLIC ACCESS EXPECTATIONS



Opportunities for public comment and engagement will be crucial to allow for iterative policy development

In Planning…

On or by **Dec 31, 2025** *NIH Revised Public Access Policy*

On or by **Dec 31, 2027** *NIH Persistent Identifiers Policy*

EASING THE ADMINISTRATIVE BURDEN FOR OPEN DATA

FACILITATING COMMUNITIES OF BEST PRACTICES

NIH's Big Picture Goals

- Work with the community to establish consistent metadata
- Develop best practices and use cases for data sharing
- Train and educate researchers on FAIR data and the importance of data sharing
- Improve discoverability of data within and across participating generalist repositories
- Facilitate greater reproducibility and reuse of data

Notice of Special Interest (NOT-OD-22-069; 01/22**)**
Supports existing data repositories to align with FAIR & TRUST principles and evaluate usage, utility, and impact

EASING THE ADMINISTRATIVE BURDEN FOR OPEN DATA

FACILITATING COMMUNITIES OF BEST PRACTICES

NIH's Big Picture Goals

- Work with the community to establish consistent metadata
- Develop best practices and use cases for data sharing
- Train and educate researchers on FAIR data and the importance of data sharing
- Improve discoverability of data within and across participating generalist repositories
- Facilitate greater reproducibility and reuse of data

Notice of Special Interest (NOT-OD-22-069; 01/22)
Supports existing data repositories to align with FAIR & TRUST principles and evaluate usage, utility, and impact

GREI Generalist Repository Ecosystem Initiative

zenodo · DRYAD · OSF · The Dataverse® Project · figshare · Mendeley Data · Vivli CENTER FOR GLOBAL CLINICAL RESEARCH DATA

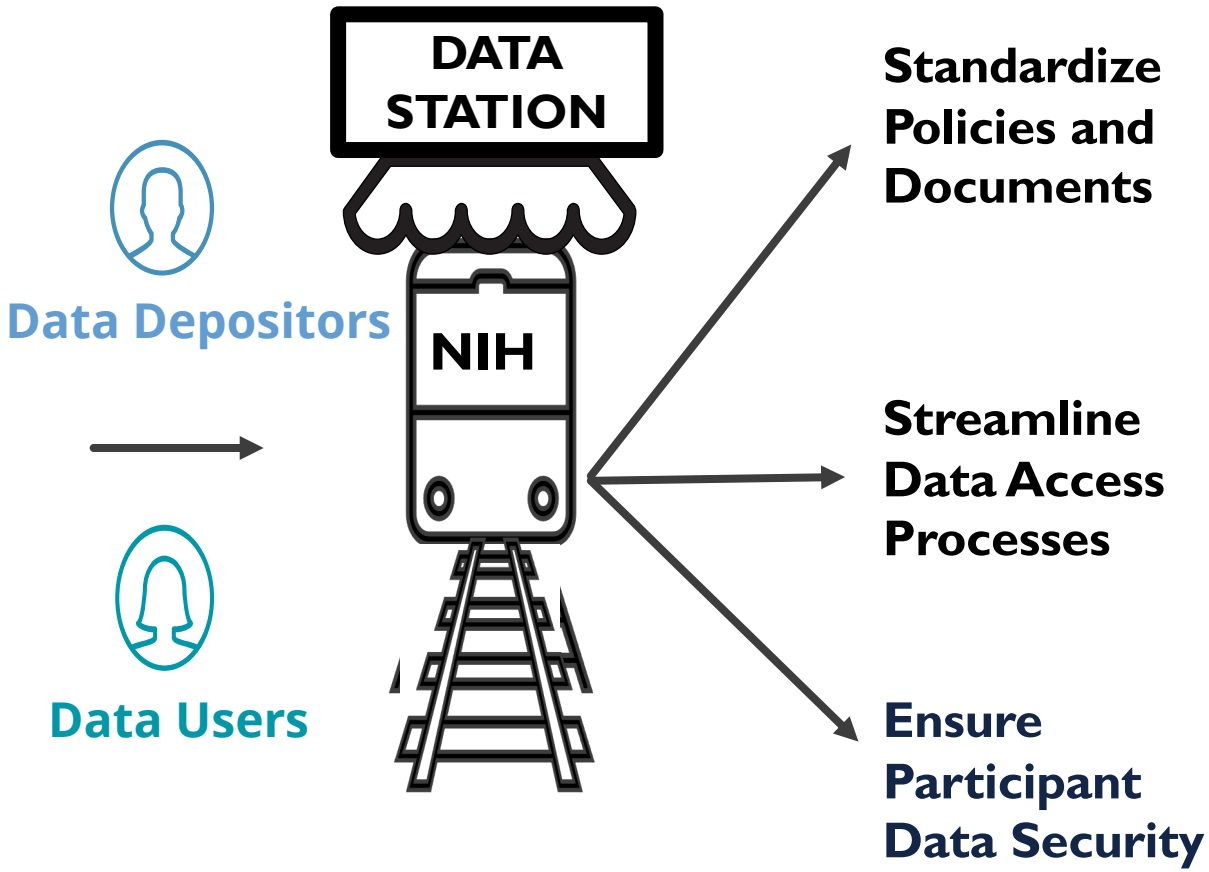
New Models of Data Stewardship

ENHANCING THE ADMINISTRATIVE GOALS FOR PARTICIPANT DATA

STREAMLINING DATA ACCESS AND MAINTAINING SECURITY

NIH's Big Picture Goals

- Develop a single policy framework that governs data access
- Develop standard language for data use agreements and related documents
- Streamline data access processes across NIH, including automating access in some cases
- Improve researchers' experiences when accessing data, including enhancing findability and tracking
- Facilitate greater efforts to protect participants from re-identification when linking data
- Facilitate responsible future use of data



STREAMLINING ACCESS AND MAINTAINING SECURITY FOR PARTICIPANT DATA

Early activities include:

- Building a comprehensive list of NIH controlled-access repositories:
  - Data submission and data access agreements used
  - Will collect terms of access
- Developing a "data agnostic" institutional certification/submission agreement
- Pilot study to increase efficiency of controlled-access data decision making
- Updating security best practices for controlled-access data repositories

BUILDING TRUST FOR THE FUTURE

- **When do people care about how their data are shared?**
  - New types of data (digital health, social media)
  - New types of analysis (algorithms, machine learning)
  - Data linkage, aggregation (for precision medicine, public health)
- **How can policymakers prioritize participant autonomy?**

LEANING ON EXTERNAL EXPERTS FOR GUIDANCE

FORECASTING DATA SCIENCE AND BIOMEDICAL RESEARCH

- **Getting beyond the Beltway…**
  - Conversations *in communities* across US
- **Building partnership and trust**
  - Webinars to prepare participants
  - Small groups, breakouts
  - Co-led and facilitated in English and Spanish
  - Following up and circling back for sustainability

ACD DISCUSSION

LOOKING TO THE FUTURE

- **What guidance/policy should NIH/federal government develop to facilitate data value and sustainability?**
- **Where should NIH build infrastructure versus invest in others?**
- **What incentives can NIH direct to promote and sustain these practices?**