Data Sharing: View from an NIH Policy Wonk

Carrie D. Wolinetz, PhD
Associate Director for Science Policy
National Institutes of Health
Roots of data sharing culture...
“The HGP changed the norms around data sharing in biomedical research.”
Data Sharing: An Essential Component
Long history…

1999
Research Tools Policy

2003
NIH Data Sharing Policy

2004
Model Organism Policy

2007
Genome-wide Association (GWAS) Policy

2008
NIH Public Access Policy (Publications)

2012
Big Data to Knowledge (BD2K) Initiative

2014
Genomic Data Sharing (GDS) Policy

White House Initiative (2013 “Holdren Memo”)

Modernization of NIH Clinical Trials
NIH Public Access Policy for Publications

• Ensures public access to published results of all research funded by NIH since 2008
  – Recipients of NIH funds required to submit final peer-reviewed journal manuscripts to PubMed Central (PMC) upon acceptance for publication.
MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren
       Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

1. Policy Principles

   The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

   Scientific research supported by the Federal Government catalyzes innovative breakthroughs that drive our economy. The results of that research become the grist for new insights and are assets for progress in areas such as health, energy, the environment, agriculture, and national security.

   Access to digital data sets resulting from federally funded research allows companies to focus resources and efforts on understanding and exploiting discoveries. For example, open weather
White House Open Data Policy

• Executive Order of May 9, 2013, Making Open and Machine Readable the New Default for Government Information
  – Requires agencies to collect or create information in a way that supports downstream information processing and dissemination activities
  – Using machine-readable and open formats, data standards, and common core and extensible metadata for all new information creation and collection efforts

• Office of Management and Budget (OMB), Sylvia Burwell, Director
  – “Open Data Policy – Managing Information as an Asset” (M-13-13)
  – Manage information as an asset throughout its life cycle to promote openness and interoperability, and properly safeguard systems and information
  – Make information resources accessible, discoverable, & usable by public
Objectives of the Holdren Memo

**Digital Data**

- Maximize free access while
- Protecting privacy and confidentiality, national security
- Recognizing intellectual property rights
- Balancing costs & benefits of long-term preservation
- Require data management plans (DMPs)
- Allow inclusion of costs in applications for funding
- Ensure appropriate evaluation of DMPs
- Monitor compliance by investigators
- Encourage deposit of data in public repositories, where possible
- Cooperate with the private sector
- Develop approaches for data citation & attribution
- Support training, education and workforce development
- Assess long-term needs for preservation and options for repositories

**Scholarly Publications**

- Public can read, download, analyze in digital form
- 12-month post-publication embargo as guideline, with stakeholder petitions to change
- Easy public search, analysis of, and access to publications
- Full public access to metadata without charge upon first publication
- Public-private collaboration
- Attribution to authors, journals, and original publishers
- Archival solutions that provide long-term preservation & access without charge
  - Uses widely available, nonproprietary standards/formats
  - Provides access for persons with disabilities (consistent with Section 508 of Rehabilitation Act)
  - Enables integration and interoperability with other Federal archival solutions and other appropriate archives
Completed Public Access Plans (as of May 2016)

Department of Agriculture (Nov 2014)

Department of Defense (Feb 2015)

Department of Energy (Jul 2014)

Department of Health & Human Services
http://www.hhs.gov/open/public-access-guiding-principles/index.html
  - Administration for Community Living [Publications] (Feb 2016)
  - Agency for Health Research & Quality (Feb 2015)
    http://www.ahrq.gov/funding/policies/publicaccess/index.html
  - Assistant Secretary for Preparedness & Response (Feb 2015)
    http://www.phe.gov/Preparedness/planning/science/Pages/AccessPlan.aspx
  - Centers for Disease Control (Jan 2015)
  - Food & Drug Administration (Feb 2015)

- National Institutes of Health (Feb 2015)

Department of Transportation (Nov 2015)
https://www.transportation.gov/open/official-dot-public-access-plan

Department of Veterans Affairs (Mar 2015)

National Institute of Standards & Technology (Apr 2015)

National Aeronautics and Space Administration (Nov 2014)
http://science.nasa.gov/media/medialibrary/2014/12/05/NASA_Plan_for_increasing_access_to_results_of_federally_funded_research.pdf

National Oceanic and Atmospheric Administration (Feb 2015)

National Science Foundation (Mar 2015)

Smithsonian Institution (Aug 2015)
http://public.media.smithsonianmag.com//file_upload_plugin/1f143b54-a9f9-4746-bef5-1c76151e3c7a.pdf

U.S. Geological Survey (Feb 2016)
http://www.usgs.gov/quality_integrity/open_access/default.asp

NOTE: Institute for Museum & Library Services and Patient-Centered Outcomes Research Institute also have policies for publications access.

See http://www.cendi.gov
NIH Plan on Digital Scientific Data

• Describes current policies and procedures and future considerations

• Maximize access by the general public, without charge, to digital scientific data

• Protect privacy, proprietary interests, and preserve the balance between the benefits of access/preservation and the costs
NIH Plan on Digital Scientific Data (cont’d)

• Explore steps to require data sharing

• Ensure that all NIH-funded researchers prepare data management and sharing plans

• Ensure that plans are reviewed during peer review

• Develop additional policies to increase public access to designated data types

• Encourage use of established repositories and community-based standards

• Develop approaches to ensure discoverability of data

• Promote interoperability and openness (M-13-13)

• Explore the development of a data commons
The results and accomplishments of the activities that NIH funds should be made available to the public. PD/PIs and recipient organizations are expected to make the results and accomplishments of their activities available to the research community and to the public at large.
From plan to policy... no easy feat?

- Data generators/data users
- What do we mean by data?
  - All data? Underlying publications?
  - Types of data? High value data?
- What so we mean by sharing?
  - FAIR principles
  - Technical details
- Complicating factors…
  - Clinical data: aggregate or IPD?
  - Avocados vs. fine wine
  - Oh, give me a home…
  - My checkbook or yours?
- Is the juice worth the squeeze?
  - Balance point between burden and value
Stay tuned…
We are not alone

  - See also *Gates Foundation*, etc.
  - ICMJE statement

- **European Medicines Agency** – Proactively releasing datasets used in marketing applications.

- **Global Alliance for Genomics and Health (GA4GH)** -- Establishing common framework of approaches to enable effective, responsible sharing of genomic and clinical data
Resources

• NIH Office of Science Policy
  – Website: http://osp.od.nih.gov/
  – Twitter: https://twitter.com/cwolinetznih
  – **Subscribe to the OSP listserv** by sending an email to LISTSERV@list.nih.gov with “Subscribe OSP_News” in the message body