NIH DATA AND RESOURCE SHARING

MPS Workshop 1:
Gauging the Impact of Requirements for Public Access to Data

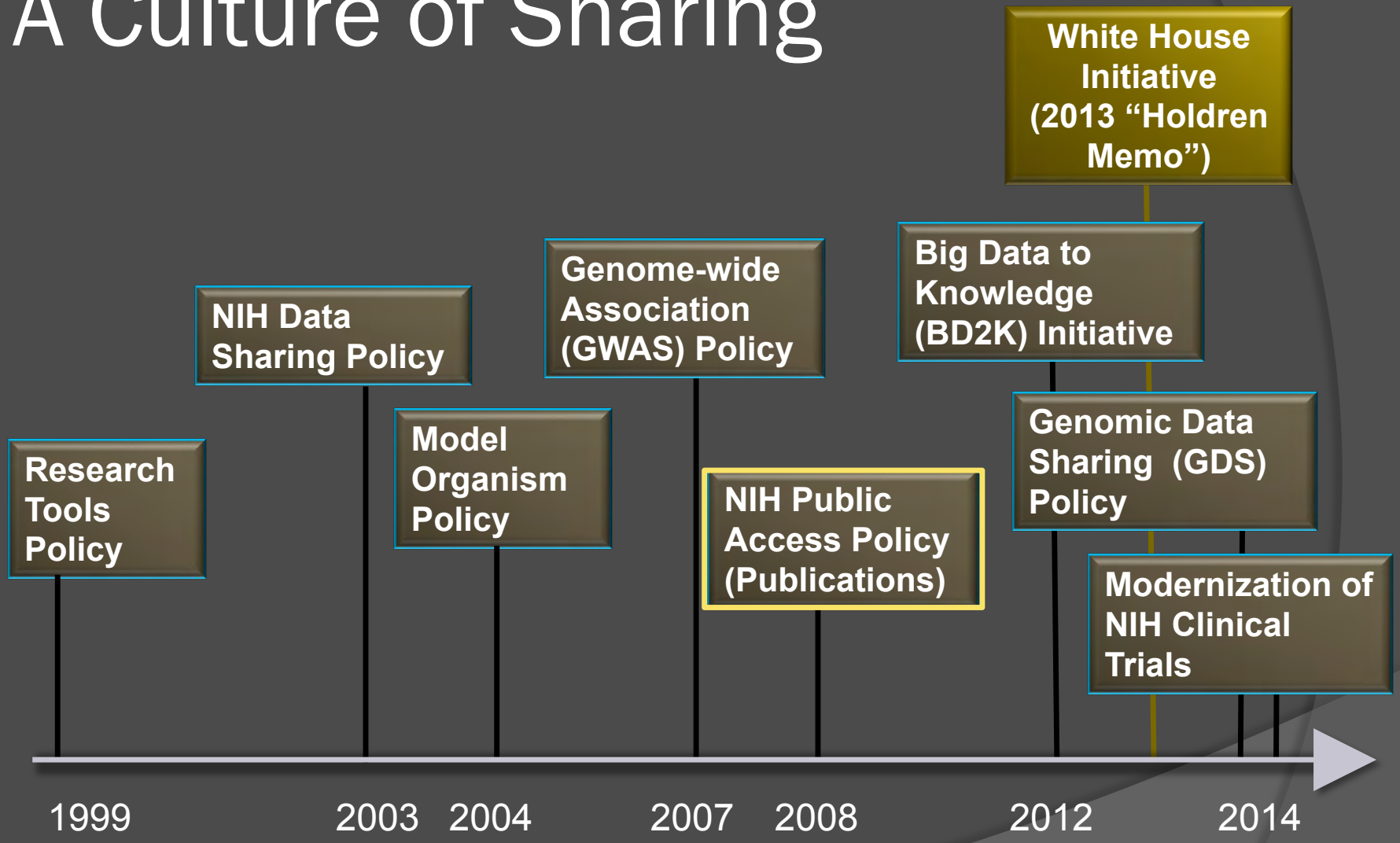
November 19, 2015

Jennie Larkin, Ph.D.
Office of the Associate Director for Data Science (ADDS), NIH

Vivien Bonazzi, Ph.D.
Office of the Associate Director for Data Science (ADDS), NIH

Office of the Director, National Institutes of Health (NIH)
U.S. Department of Health & Human Services (HHS)

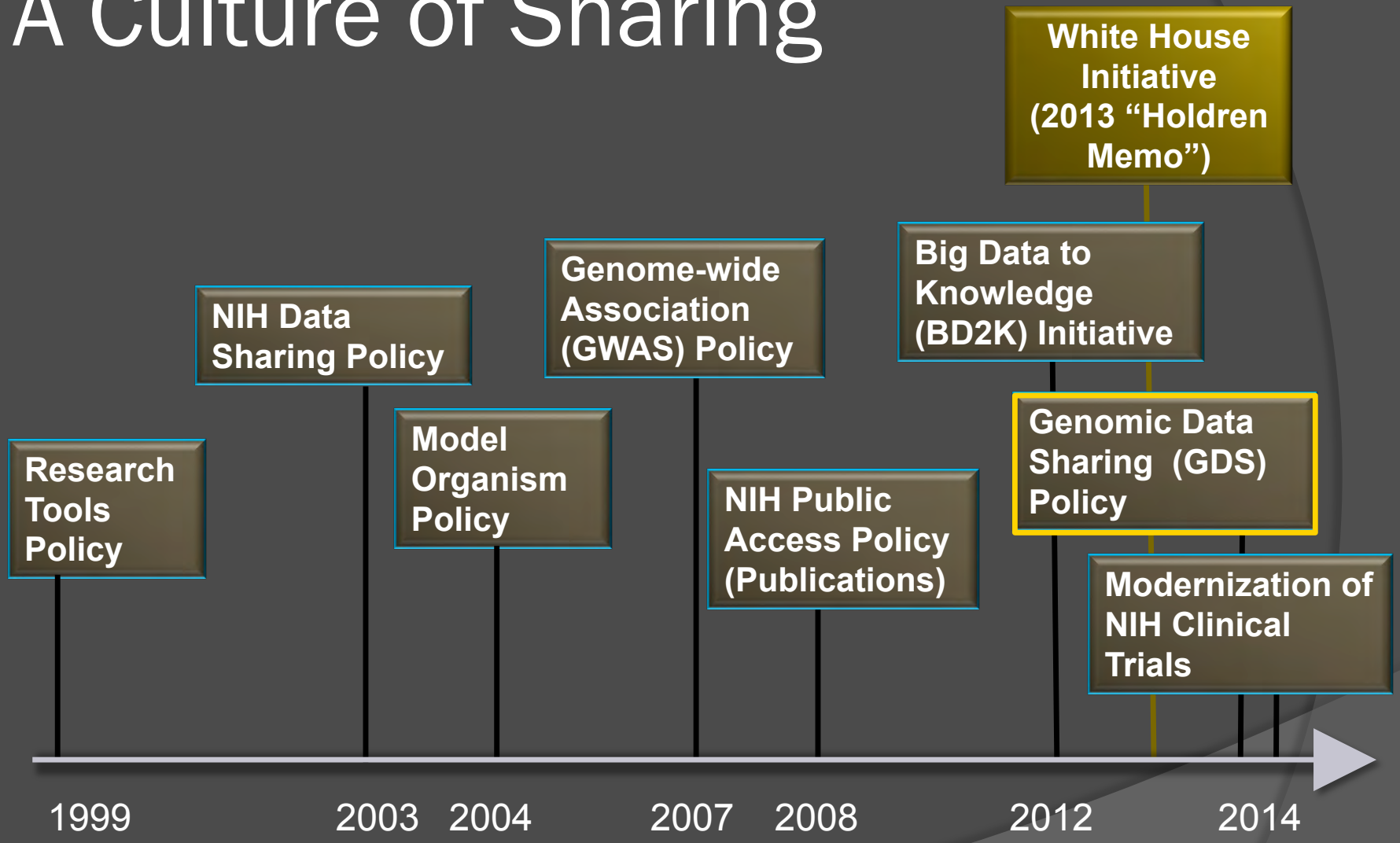
A Culture of Sharing



NIH Public Access Policy for Publications

- Since 2005, public access to published results of all research funded by NIH (by law since 2008)
  - Recipients of NIH funds required to submit final peer-reviewed journal manuscripts to PubMed Central (PMC) upon acceptance for publication
  - Papers must be accessible to the public on PMC no later than 12 months after publication

A Culture of Sharing



Purpose and Scope of the NIH Genomic Data Sharing (GDS) Policy

- **Purpose**
  - Sets forth expectations and responsibilities that ensure the broad and responsible sharing of genomic research data in a timely manner

- **Scope**
  - All NIH-funded research generating large-scale **human** or **non-human genomic data** and the use of these data for subsequent research
    - Data to be submitted to NIH-designated data repositories (e.g., dbGaP, GEO, GenBank, WormBase, FlyBase, Rat Genome Database)
  - Applies to all funding mechanisms (grants, contracts, intramural support) – Effective date was January 25, 2015
  - No minimum threshold for cost

ENCODE

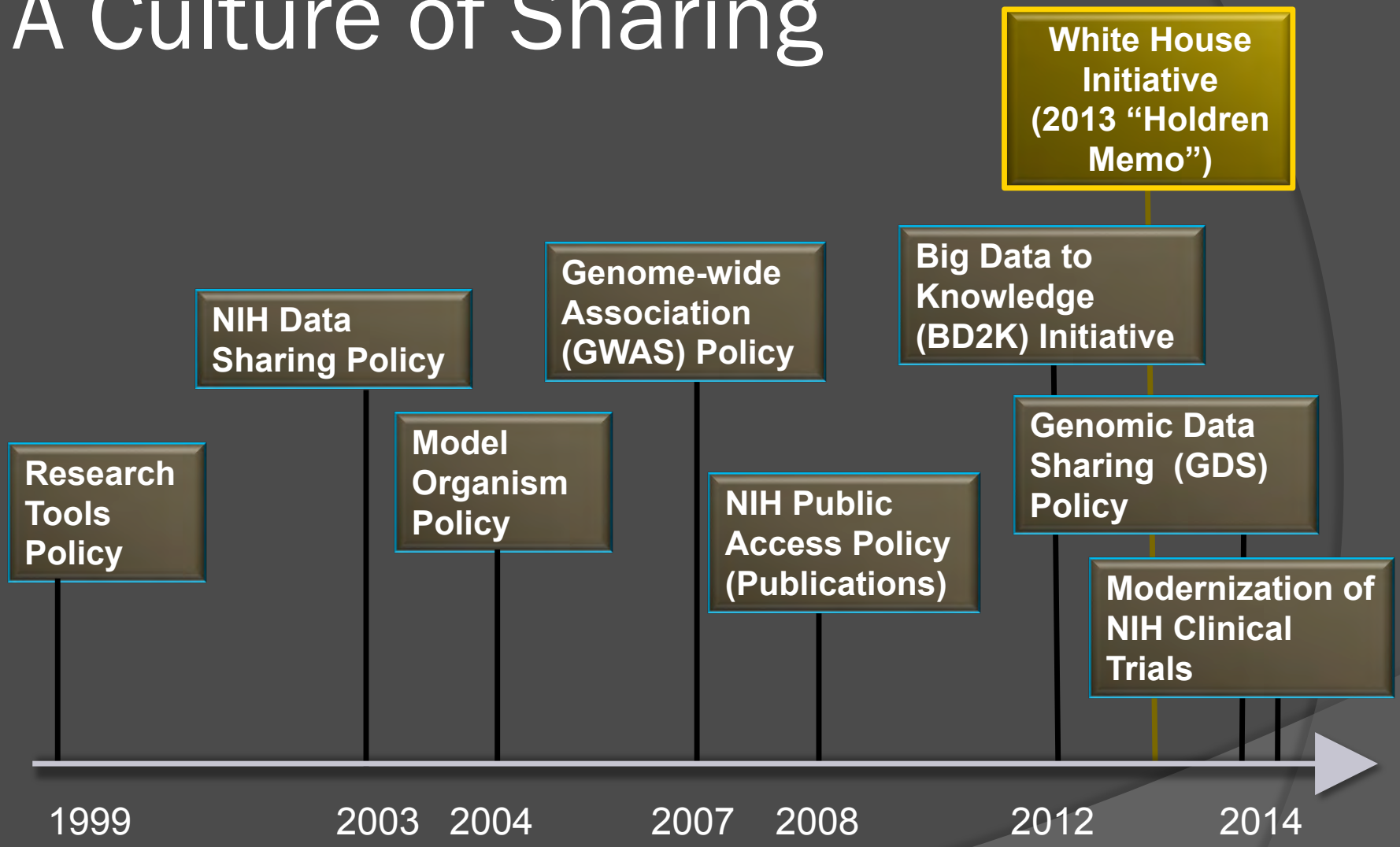
International HapMap Project

THE CANCER GENOME ATLAS

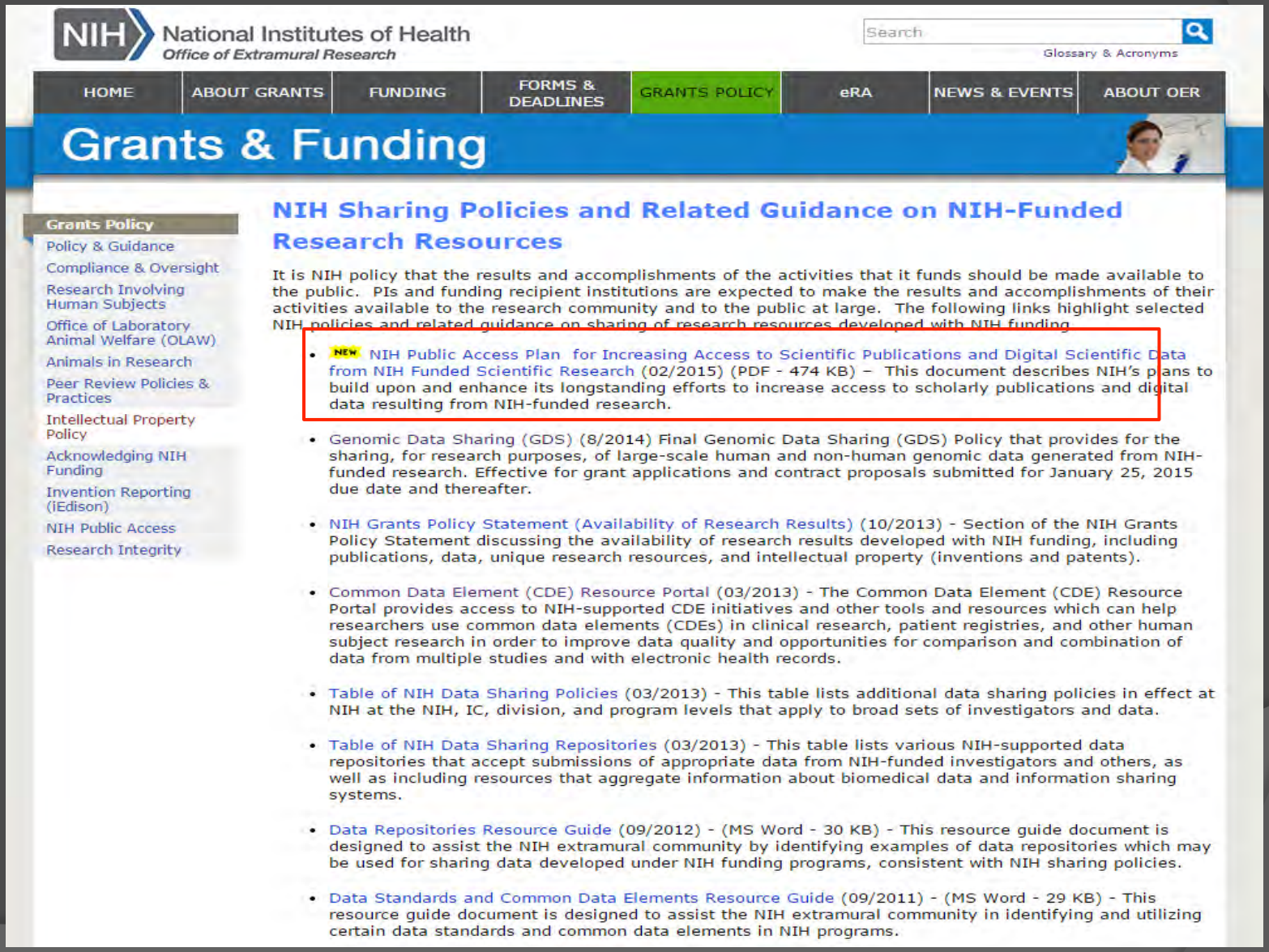
Data Sharing: An Essential Component

1000 Genomes
A Deep Catalog of Human Genetic Variation

A Culture of Sharing



NIH Response to Administration Directives



NIH National Institutes of Health
Office of Extramural Research

Search
Glossary & Acronyms

HOME | ABOUT GRANTS | FUNDING | FORMS & DEADLINES | GRANTS POLICY | eRA | NEWS & EVENTS | ABOUT OER

Grants & Funding

Grants Policy
- Policy & Guidance
- Compliance & Oversight
- Research Involving Human Subjects
- Office of Laboratory Animal Welfare (OLAW)
- Animals in Research
- Peer Review Policies & Practices
- Intellectual Property Policy
- Acknowledging NIH Funding
- Invention Reporting (iEdison)
- NIH Public Access
- Research Integrity

NIH Sharing Policies and Related Guidance on NIH-Funded Research Resources

It is NIH policy that the results and accomplishments of the activities that it funds should be made available to the public. PIs and funding recipient institutions are expected to make the results and accomplishments of their activities available to the research community and to the public at large. The following links highlight selected NIH policies and related guidance on sharing of research resources developed with NIH funding.

- NEW NIH Public Access Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research (02/2015) (PDF - 474 KB) – This document describes NIH's plans to build upon and enhance its longstanding efforts to increase access to scholarly publications and digital data resulting from NIH-funded research.
- Genomic Data Sharing (GDS) (8/2014) Final Genomic Data Sharing (GDS) Policy that provides for the sharing, for research purposes, of large-scale human and non-human genomic data generated from NIH-funded research. Effective for grant applications and contract proposals submitted for January 25, 2015 due date and thereafter.
- NIH Grants Policy Statement (Availability of Research Results) (10/2013) - Section of the NIH Grants Policy Statement discussing the availability of research results developed with NIH funding, including publications, data, unique research resources, and intellectual property (inventions and patents).
- Common Data Element (CDE) Resource Portal (03/2013) - The Common Data Element (CDE) Resource Portal provides access to NIH-supported CDE initiatives and other tools and resources which can help researchers use common data elements (CDEs) in clinical research, patient registries, and other human subject research in order to improve data quality and opportunities for comparison and combination of data from multiple studies and with electronic health records.
- Table of NIH Data Sharing Policies (03/2013) - This table lists additional data sharing policies in effect at NIH at the NIH, IC, division, and program levels that apply to broad sets of investigators and data.
- Table of NIH Data Sharing Repositories (03/2013) - This table lists various NIH-supported data repositories that accept submissions of appropriate data from NIH-funded investigators and others, as well as including resources that aggregate information about biomedical data and information sharing systems.
- Data Repositories Resource Guide (09/2012) - (MS Word - 30 KB) - This resource guide document is designed to assist the NIH extramural community by identifying examples of data repositories which may be used for sharing data developed under NIH funding programs, consistent with NIH sharing policies.
- Data Standards and Common Data Elements Resource Guide (09/2011) - (MS Word - 29 KB) - This resource guide document is designed to assist the NIH extramural community in identifying and utilizing certain data standards and common data elements in NIH programs.

NIH Plan on 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 (continued)

- Ensure that all NIH-funded researchers prepare data management and sharing plans
- Ensure that plans are reviewed during peer review
- Encourage use of established repositories and community-based standards

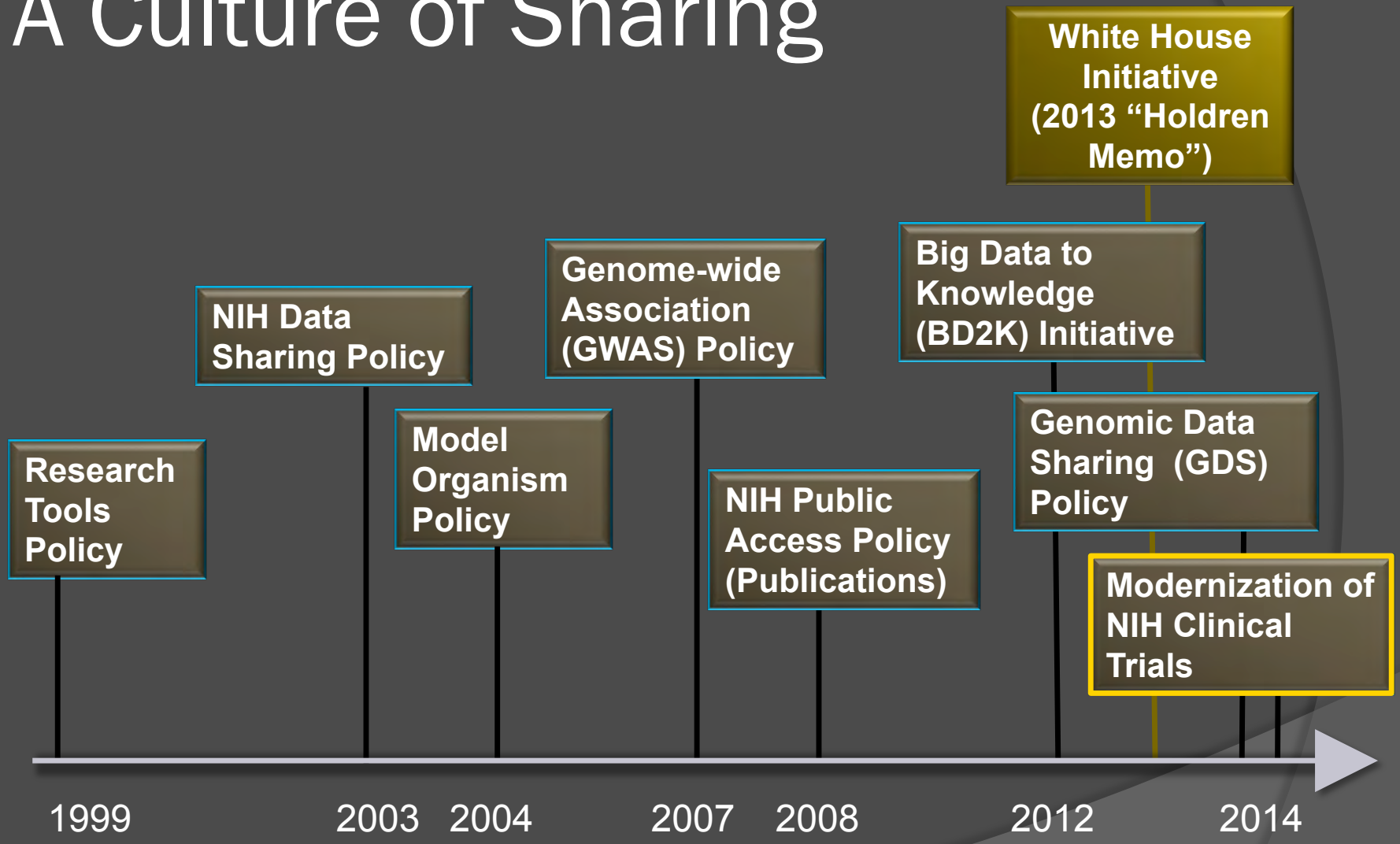
NIH Plan on Digital Scientific Data (continued)

- Develop approaches to ensure discoverability of data
- Implement the data commons

Considerations for Implementation

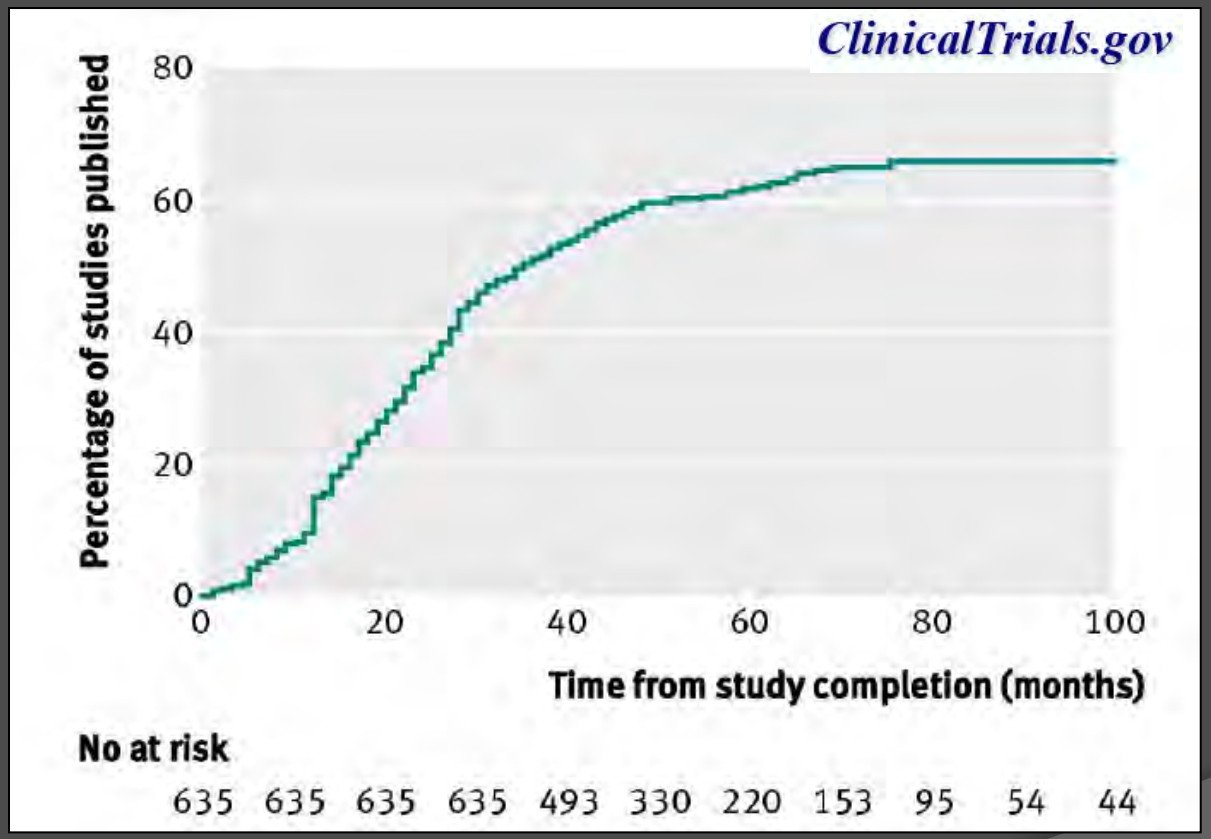
- Timelines for implementation
- For human data: privacy, confidentiality, informed consent issues
- Developing NIH-wide approaches for sustaining data resources
- Implementation costs within the constraints of existing budgets and resources

A Culture of Sharing



Modernizing NIH Clinical Trials Activities: The Need

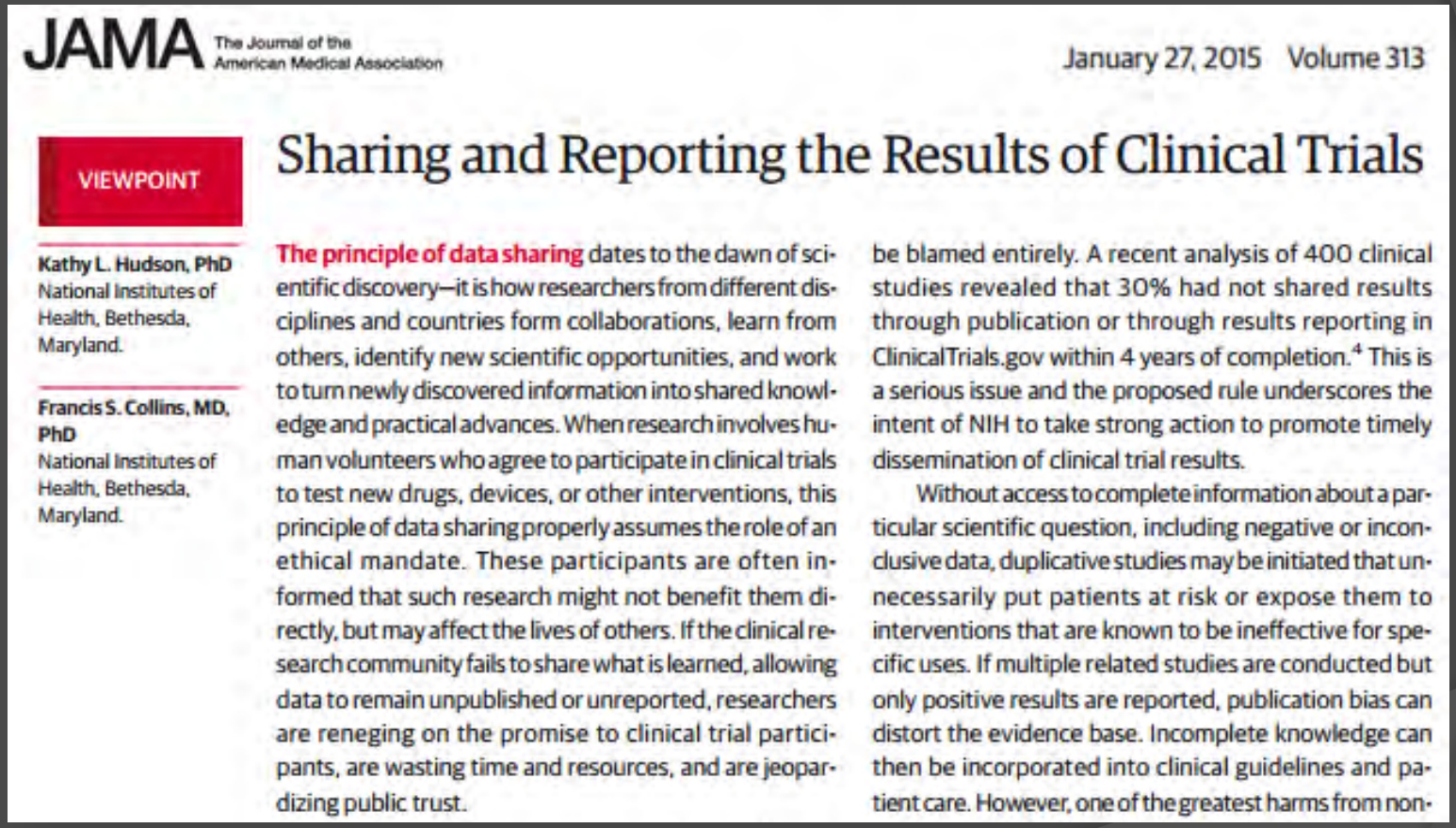
- NIH-Funded trials published within 100 months of completion



Less than 50% published within 30 months of completion

BMJ 2012;344:d7292

Modernizing NIH Clinical Trials Activities: Call to Action



JAMA The Journal of the American Medical Association

January 27, 2015 Volume 313

VIEWPOINT

Sharing and Reporting the Results of Clinical Trials

Kathy L. Hudson, PhD
National Institutes of Health, Bethesda, Maryland.

Francis S. Collins, MD, PhD
National Institutes of Health, Bethesda, Maryland.

The principle of data sharing dates to the dawn of scientific discovery—it is how researchers from different disciplines and countries form collaborations, learn from others, identify new scientific opportunities, and work to turn newly discovered information into shared knowledge and practical advances. When research involves human volunteers who agree to participate in clinical trials to test new drugs, devices, or other interventions, this principle of data sharing properly assumes the role of an ethical mandate. These participants are often informed that such research might not benefit them directly, but may affect the lives of others. If the clinical research community fails to share what is learned, allowing data to remain unpublished or unreported, researchers are reneging on the promise to clinical trial participants, are wasting time and resources, and are jeopardizing public trust.

be blamed entirely. A recent analysis of 400 clinical studies revealed that 30% had not shared results through publication or through results reporting in ClinicalTrials.gov within 4 years of completion.[4] This is a serious issue and the proposed rule underscores the intent of NIH to take strong action to promote timely dissemination of clinical trial results.

Without access to complete information about a particular scientific question, including negative or inconclusive data, duplicative studies may be initiated that unnecessarily put patients at risk or expose them to interventions that are known to be ineffective for specific uses. If multiple related studies are conducted but only positive results are reported, publication bias can distort the evidence base. Incomplete knowledge can then be incorporated into clinical guidelines and patient care. However, one of the greatest harms from non-

Common Data Elements (CDEs)

- Encouraging knowledge and adoption of CDEs in NIH-sponsored clinical research.
- NIH CDE Workshop (September 30, 2015)
- Facilitate harmonization of platforms, tools, and resources across NIH Institutes and Centers.
  - NCI caDSR: https://wiki.nci.nih.gov/display/caDSR/caDSR+CDE+Browser
  - BMIC CDE Portal: https://www.nlm.nih.gov/cde/

BRICS | Biomedical Research Informatics Computing System

FDAAA and Clinical Trial reporting

- FDAAA (2007) mandates results sharing through ClinicalTrials.gov for most FDA-regulated studies
- NPRM and NIH Policy expanding this scope (2014 – 2016)

National Institutes of Health
Turning Discovery Into Health

THE *COMMONS*

What are the PRINCIPLES of *The Commons*?

- Crucial component of a digital biomedical ecosystem
- Treats products of research – data, software, methods, papers etc. as *digital research objects*
- Digital research objects exist in a shared virtual space *Find, Deposit, Manage, Share and Reuse data, software, metadata and workflows*
- Digital objects need to conform to FAIR principles:
  - **F**indable
  - **A**ccessible (*and usable)*
  - **I**nteroperable
  - **R**eusable

What is The *Commons Framework*?

- Exploits new scalable computing technologies - Cloud
- Provides physical or logical access to data
- *Simplifies* access, sharing and interoperability of digital research objects such as data, software, metadata and workflows
- Makes digital research objects indexable and findable: FAIR
- Provides understanding and accounting of usage patterns
- Is potentially more *cost effective* given digital growth
- Gives currency to digital objects and the people who develop and support them

The Commons Framework

SaaS

PaaS

IaaS

Software: Services & Tools

App store/User Interface

scientific analysis tools/workflows

Services: APIs, Containers, Indexing,

Data

"Reference" Data Sets

User defined data

Compute Platform: Cloud or HPC

Digital Object Compliance

Commons: *Digital Object Compliance*

- **Attributes of digital research objects in the Commons**

Initial Phase

- Unique digital object identifiers resolvable to original authoritative source
- Machine readable
- A minimal set of searchable metadata
- Physically available in a cloud-based Commons provider
- Clear access rules (especially important for human subjects data)
- An entry (with metadata) in one or more indices

Future Phases

- Standard, community-based unique digital object identifiers
- Conform to community-approved standard metadata for enhanced searching
- Digital objects accessible via open standard APIs
- Are physically and logical available to the Commons

International Efforts and Activities --Growing

- *Wellcome Trust* -- Publications and datasets shared http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtp053977.pdf.
  - See also *Gates Foundation*, etc.
- *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

Thank you!

NIH Sharing Policies Website: http://sharing.nih.gov

Email: Sharing@nih.gov

NIH Office of the Associate Director for Data Science (ADDS)

- *Data Science Blog – Input | Output*
  https://datascience.nih.gov/blog

NIH Office of Extramural Research (OER)

- *OER Blog – Open Mike*
  http://nexus.od.nih.gov/all/category/open-mike/

NIH Office of Science Policy (OSP)

- *OSP Blog - Under the Poliscope:*
  http://osp.od.nih.gov/under-the-poliscope