Developing a New Culture: NIH Policies in Data Management, Sharing & Access

Jaime M. Guidry Auvil, Ph.D.

Office of Data Sharing, NCI, NIH

NIH NATIONAL CANCER INSTITUTE

January 18, 2023

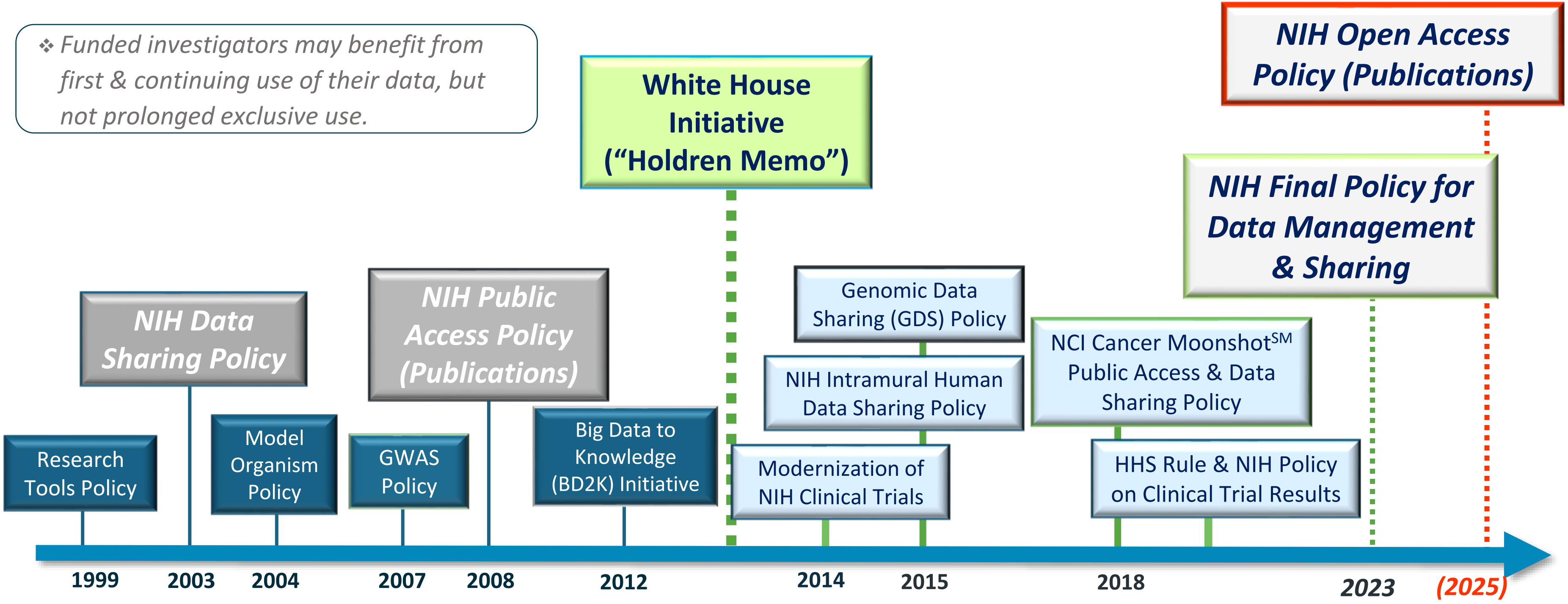
What is “Open Science”?

A “movement” to make scientific research (including ***publications, DATA, physical samples, and software***) and dissemination accessible to all levels of society, amateur or professional.

- Open science is transparent and accessible knowledge that is shared and developed through **collaborative networks**.
- It encompasses practices such as:
  - publishing open research & campaigning for **open access**,
  - encouraging scientists to practice open-notebook science (such as **openly sharing data and code**),
  - broader dissemination and engagement in science, and
  - generally making it easier to publish, access and communicate scientific knowledge.
- Usage of the term varies substantially across disciplines, with a notable prevalence in the STEM disciplines.

NIH History of Data Sharing Policies

- *Funded investigators may benefit from first & continuing use of their data, but not prolonged exclusive use.*



Research Tools Policy — 1999

NIH Data Sharing Policy — 2003

Model Organism Policy — 2004

GWAS Policy — 2007

NIH Public Access Policy (Publications) — 2008

Big Data to Knowledge (BD2K) Initiative — 2012

White House Initiative (“Holdren Memo”)

Modernization of NIH Clinical Trials — 2014

Genomic Data Sharing (GDS) Policy

NIH Intramural Human Data Sharing Policy — 2015

NCI Cancer Moonshot^{SM} Public Access & Data Sharing Policy

HHS Rule & NIH Policy on Clinical Trial Results — 2018

NIH Final Policy for Data Management & Sharing — *2023*

NIH Open Access Policy (Publications) — *(2025)*

Investigators must share any information necessary to understand, develop or reproduce published research (raw data, statistical methods, tools, source code)

Key Messages for Final NIH DMS Policy

Promote ***open science***, stimulate new ***discovery***, enable ***rigor*** & ***reproducibility***, and provide ***transparency***

Driving A Cultural Shift through planning for consistent, collaborative & impactful data management and sharing as a critical part of all research

NIH is taking a "***learning approach***" (i.e., phased and iterative implementation in the years to come).

Over time, thoughtful DMS plans will inform clear guidance on the highest value data types beyond genomics (repository, timelines, etc.)

NIH is providing resources, training, and guidance, to the extent possible, for initial rollout & will continue to develop these over time

Final NIH Policy for Data Management & Sharing (DMS)

Goal: Share scientific data supported by the NIH broadly and immediately

Scope

- ***All NIH-supported research that generates scientific data***
  - Intramural & Extramural research
  - Grants, Contracts & OTAs
  - Human & non-human data; no budget threshold
- Submitted on/after ***Jan. 25, 2023***
- Complementary to current data sharing policies (does not replace)

Expectations

✓ **Investigators must submit *plans to manage & share data (research applications)***

- PIs should comply with app
DMS Plans, and share data
  - In set repositories, to extent
possible ***(FAIR Principles)***
  - At publication or end of awa
whichever sooner

Public Access

- Final manuscripts to PubMed Central w/i 12 months of acceptance

****Public Access policies to be updated asap (by Dec. 31, 2025)*** → *agencies make publications & supporting data from federally funded research publicly accessible with no embargo on their free & broad release*

DMS Policy Applies to All *Scientific Data*

“The recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications."

Not Scientific Data Under the DMS Policy:

- *Data not necessary for, or of sufficient quality, to validate & replicate research findings*
- *Laboratory notebooks*
- *Preliminary analyses*
- *Completed case report forms*
- *Drafts of scientific papers,*
- *Plans for future research*
- *Peer reviews*
- *Communications with colleagues, or*
- *Physical objects, (e.g., laboratory specimens)*

Responsibilities and Expectations: Investigators

Researchers to prospectively plan for how scientific data will be managed and shared through submission of a DMS Plan that considers any potential restrictions or limitations (no separate GDS Plans)

Researchers to maintain alignment with the approved DMS Plans by the NIH ICO (in the awards) and with the evolution of Plans during project period (through RPPR)

NIH DMS: Supplemental Guidance

Recommended Elements of a NIH DMS Plan

Data Types

Tools for Access/ Manipulation

Data Access/Reuse Considerations

Data Standards

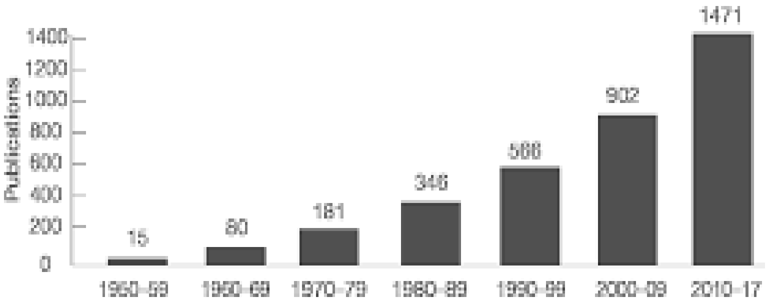
Sharing Oversight Methods

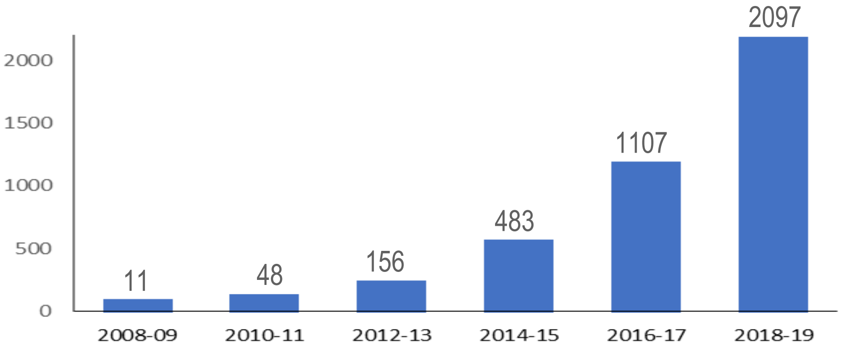
Data Access & Timelines

Guidance for Selecting a Repository for Data

Guidance for Allowable Costs of Data Management & Sharing

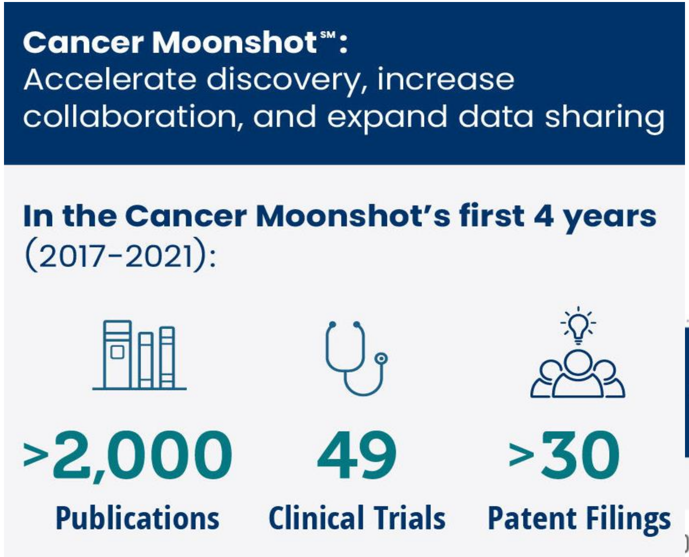
Driving Science through Publications, Data & Collaboration





	Framingham Heart Study	The Cancer Genome Atlas
Study Length	70 years	12 years
Cases Studied	15,144	11,429
Publications	**3,698 (~38,000 PMC)**	**3,747 (~62,000 PMC)**
Controlled-access Data	Consortia; HMB (+IRB/MDS, 2K=NPU)	Collaborative Teams & Public Use of Data; GRU
Authorized Users	715	3,335
Open Data Use & Availability Timing	Little Open Data; mostly available with publication	Some Open Data; All data immediately available to community

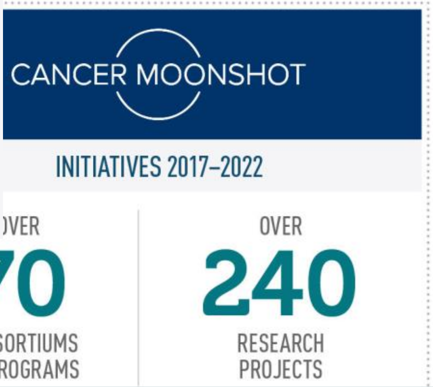
The Cancer Moonshot: Success in Mission-Driven Science



Cancer Moonshot℠:
Accelerate discovery, increase collaboration, and expand data sharing

In the Cancer Moonshot's first 4 years (2017–2021):

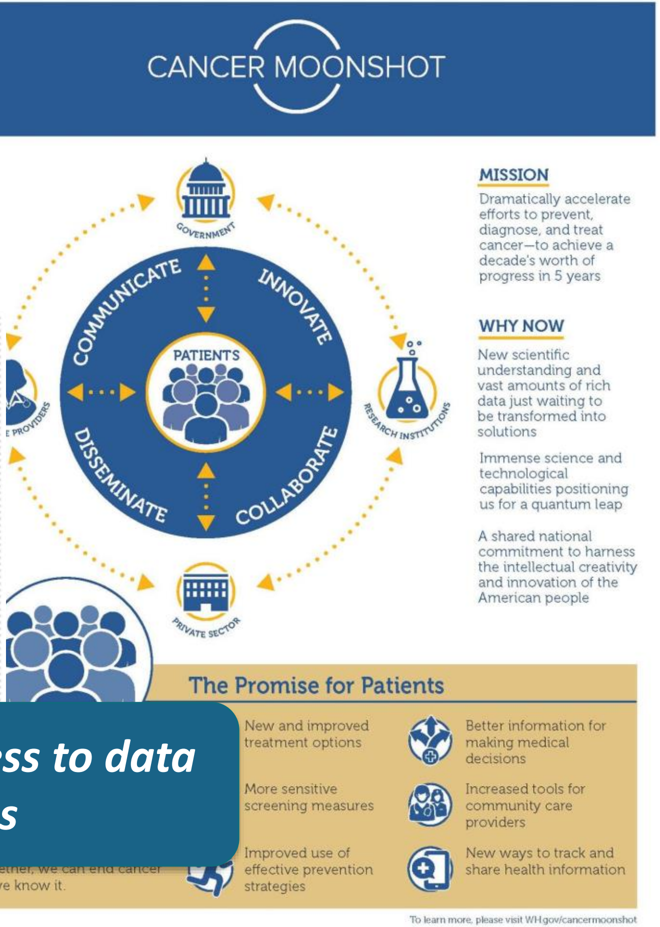
>2,000	49	>30
Publications	Clinical Trials	Patent Filings



CANCER MOONSHOT

INITIATIVES 2017–2022

OVER	OVER
70	240
CONSORTIUMS OR PROGRAMS	RESEARCH PROJECTS



CANCER MOONSHOT

MISSION

Dramatically accelerate efforts to prevent, diagnose, and treat cancer—to achieve a decade's worth of progress in 5 years

WHY NOW

New scientific understanding and vast amounts of rich data just waiting to be transformed into solutions

Immense science and technological capabilities positioning us for a quantum leap

A shared national commitment to harness the intellectual creativity and innovation of the American people

The Promise for Patients

New and improved treatment options

Better information for making medical decisions

More sensitive screening measures

Increased tools for community care providers

Improved use of effective prevention strategies

New ways to track and share health information

Together, we can end cancer as we know it.

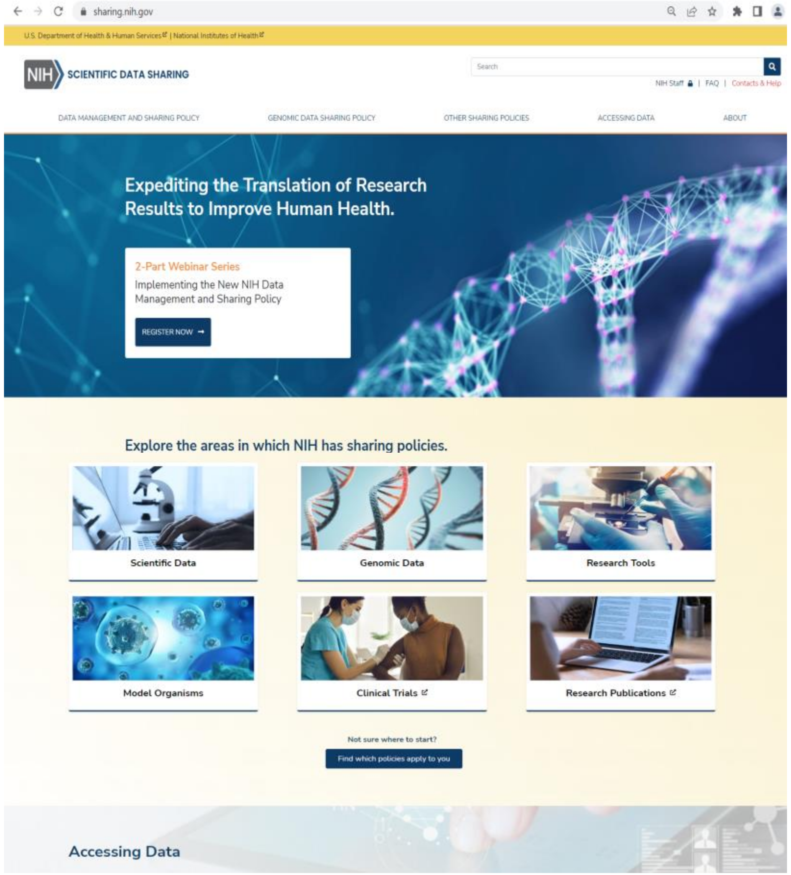
To learn more, please visit WH.gov/cancermoonshot

*****Take Home Message: purposeful, broad, early access to data leads to much faster and impactful outcomes***

NIH Resources to Support Investigators

https://sharing.nih.gov

- **A Central Data Sharing Website**: A one-stop-shop for information on NIH data sharing policies & related resources
  - *NIH Policy Notices & Supplemental Guidance*
  - *Relevant Forms & Templates*
  - *Frequently Asked Questions (FAQs)*
  - *Sample DMS Plans*
  - *Policy Decision Tool*
  - *Links to Individual NIH Institutes & Centers for program-specific references*
- **Other Resources and Tools** [e.g., DMPTool (https://dmptool.org/), Repositories]
- **NIH Training**: News Events
- **A Central Mailbox:** Help answer DMS questions (Sharing@nih.mail.gov).



sharing.nih.gov

U.S. Department of Health & Human Services | National Institutes of Health

NIH SCIENTIFIC DATA SHARING

Search
NIH Staff | FAQ | Contacts & Help
DATA MANAGEMENT AND SHARING POLICY
GENOMIC DATA SHARING POLICY
OTHER SHARING POLICIES
ACCESSING DATA
ABOUT

Expediting the Translation of Research Results to Improve Human Health.

2-Part Webinar Series

Implementing the New NIH Data Management and Sharing Policy

Explore the areas in which NIH has sharing policies.

Not sure where to start?

Accessing Data

NATIONAL
CANCER
INSTITUTE

www.cancer.gov

www.cancer.gov/espanol

Additional Slides

Benefits of Broad Data Sharing

Collaborator Sharing

- Between investigator to investigator (e.g., sharing upon publication and request to the author)

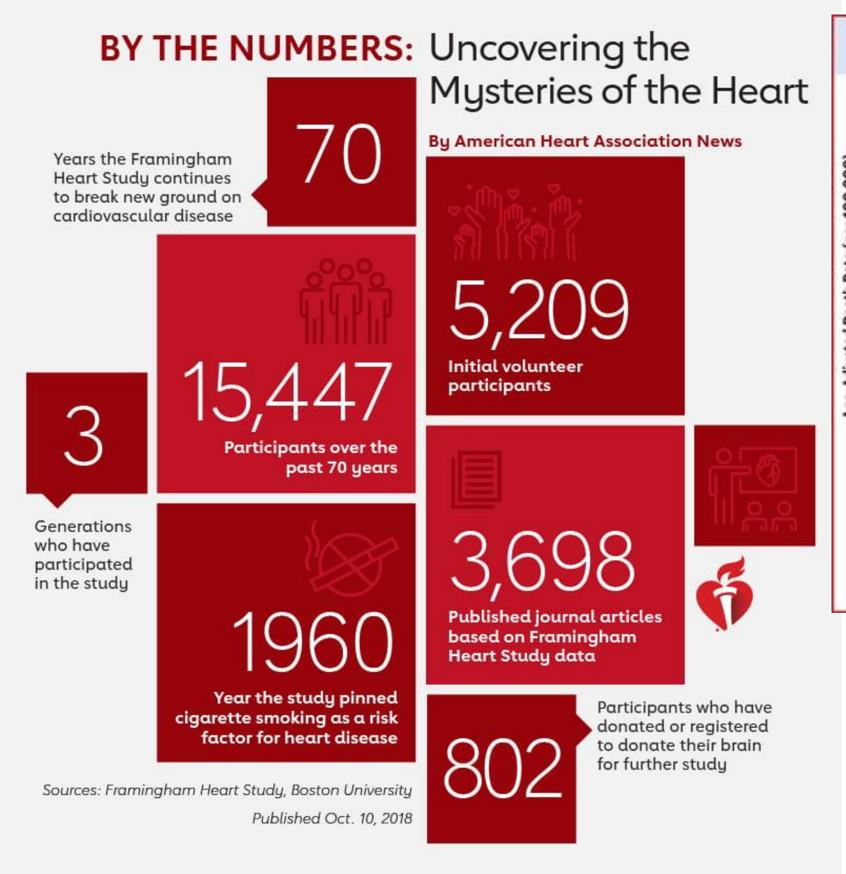
Consortium Sharing

- Within large collaborative groups (e.g., sharing between investigators within a consortium/ network)

Broad Sharing

- Ensures fair and equitable access and secondary use of data by the wider research community (e.g., NIH's Genomic Data Sharing Policy)
- Has the most impact on *driving scientific innovation and discovery and ensuring replication of results*
- Broad sharing ≠ Open access data

Framingham Study: Success in Data Collection Over Time



BY THE NUMBERS: Uncovering the Mysteries of the Heart

By American Heart Association News

70 Years the Framingham Heart Study continues to break new ground on cardiovascular disease

5,209 Initial volunteer participants

15,447 Participants over the past 70 years

3 Generations who have participated in the study

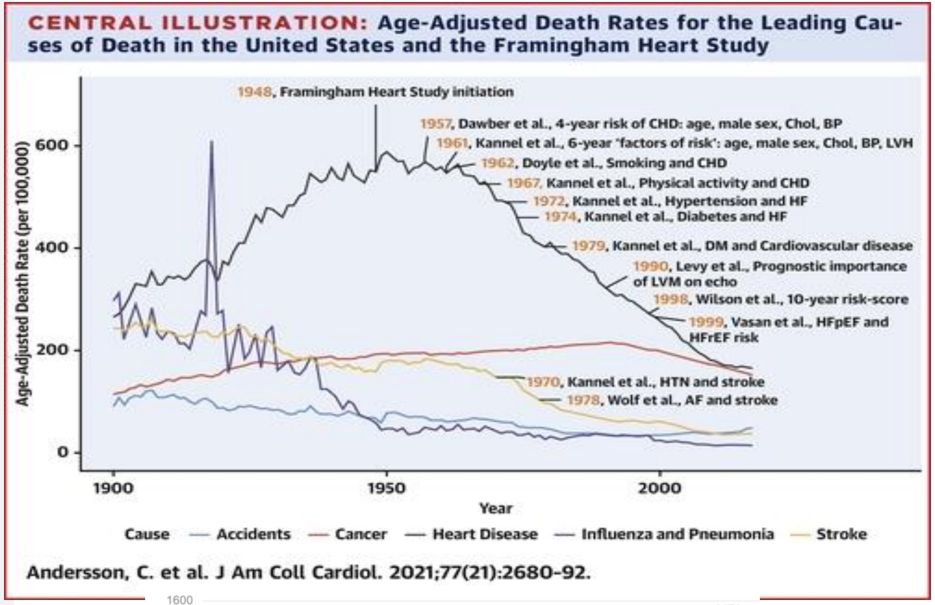
3,698 Published journal articles based on Framingham Heart Study data

1960 Year the study pinned cigarette smoking as a risk factor for heart disease

802 Participants who have donated or registered to donate their brain for further study

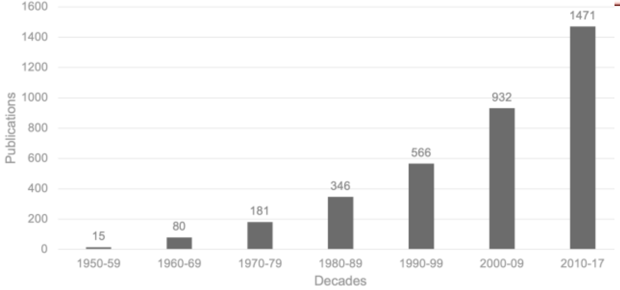
Sources: Framingham Heart Study, Boston University

Published Oct. 10, 2018



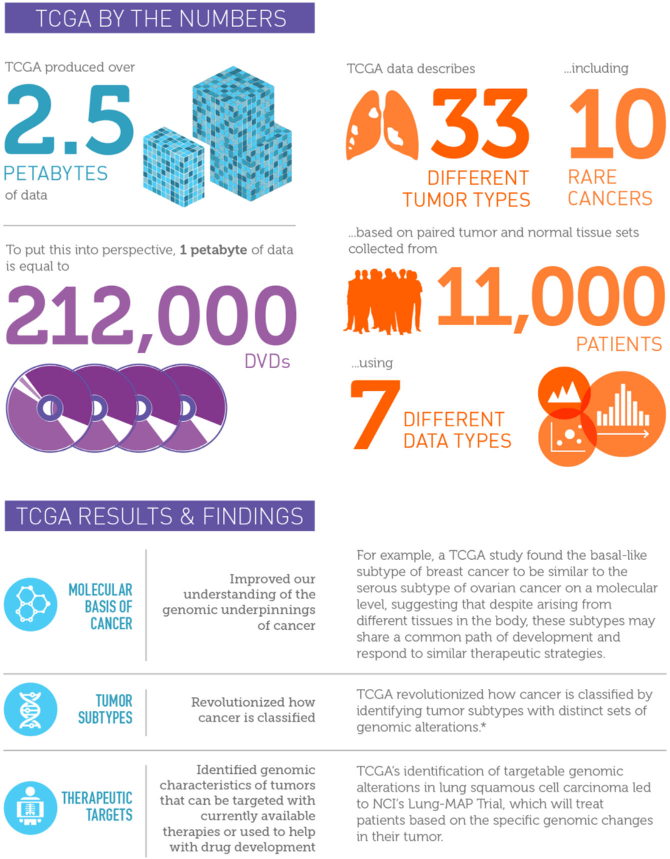
CENTRAL ILLUSTRATION: Age-Adjusted Death Rates for the Leading Causes of Death in the United States and the Framingham Heart Study

Andersson, C. et al. J Am Coll Cardiol. 2021;77(21):2680-92.



Total articles published through November 2017 = 3,561

The Cancer Genome Atlas: Success in Open Team Science



TCGA BY THE NUMBERS

TCGA produced over **2.5** PETABYTES of data

To put this into perspective, **1 petabyte** of data is equal to **212,000** DVDs

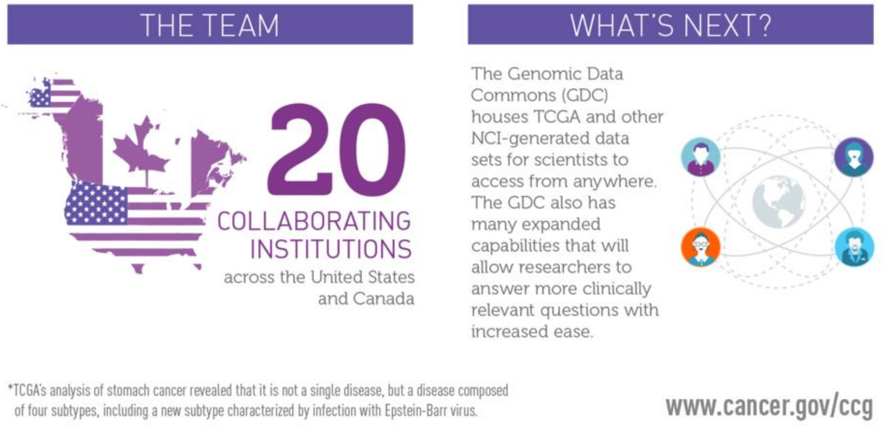
TCGA data describes **33** DIFFERENT TUMOR TYPES ...including **10** RARE CANCERS

...based on paired tumor and normal tissue sets collected from **11,000** PATIENTS

...using **7** DIFFERENT DATA TYPES

TCGA RESULTS & FINDINGS

MOLECULAR BASIS OF CANCER	Improved our understanding of the genomic underpinnings of cancer	For example, a TCGA study found the basal-like subtype of breast cancer to be similar to the serous subtype of ovarian cancer on a molecular level, suggesting that despite arising from different tissues in the body, these subtypes may share a common path of development and respond to similar therapeutic strategies.
TUMOR SUBTYPES	Revolutionized how cancer is classified	TCGA revolutionized how cancer is classified by identifying tumor subtypes with distinct sets of genomic alterations.*
THERAPEUTIC TARGETS	Identified genomic characteristics of tumors that can be targeted with currently available therapies or used to help with drug development	TCGA's identification of targetable genomic alterations in lung squamous cell carcinoma led to NCI's Lung-MAP Trial, which will treat patients based on the specific genomic changes in their tumor.



THE TEAM

20 COLLABORATING INSTITUTIONS across the United States and Canada

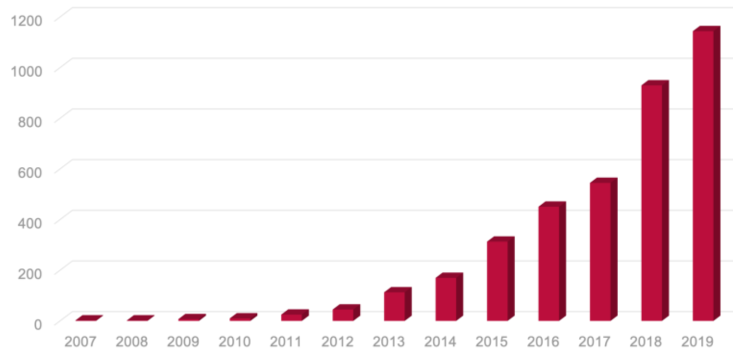
WHAT'S NEXT?

The Genomic Data Commons (GDC) houses TCGA and other NCI-generated data sets for scientists to access from anywhere. The GDC also has many expanded capabilities that will allow researchers to answer more clinically relevant questions with increased ease.

*TCGA's analysis of stomach cancer revealed that it is not a single disease, but a disease composed of four subtypes, including a new subtype characterized by infection with Epstein-Barr virus.

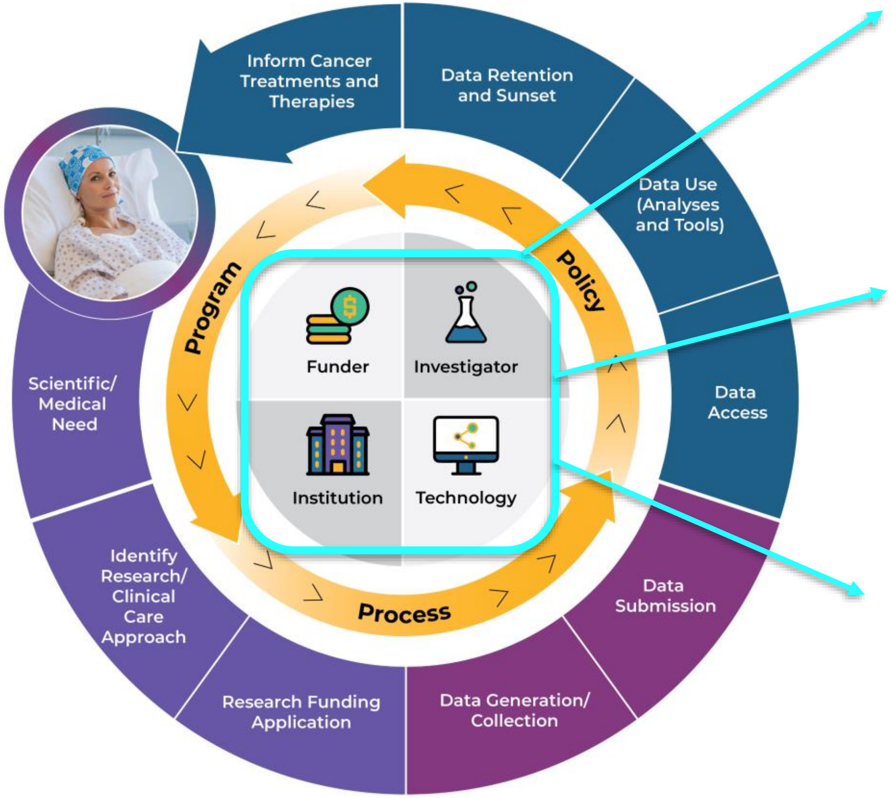
www.cancer.gov/ccg

Number of Publications Using TCGA Data



NIH NATIONAL CANCER INSTITUTE

Scientific Data Lifecycle: Keys to Impactful Discovery


Inform Cancer Treatments and Therapies
Data Retention and Sunset
Data Use (Analyses and Tools)
Data Access
Data Submission
Data Generation/ Collection
Research Funding Application
Identify Research/ Clinical Care Approach
Scientific/ Medical Need
Program
Policy
Process
Funder
Investigator
Institution
Technology

Critical Questions to Answer

Programs that define therapeutic needs and essential scientific gaps to be filled using structured datasets.

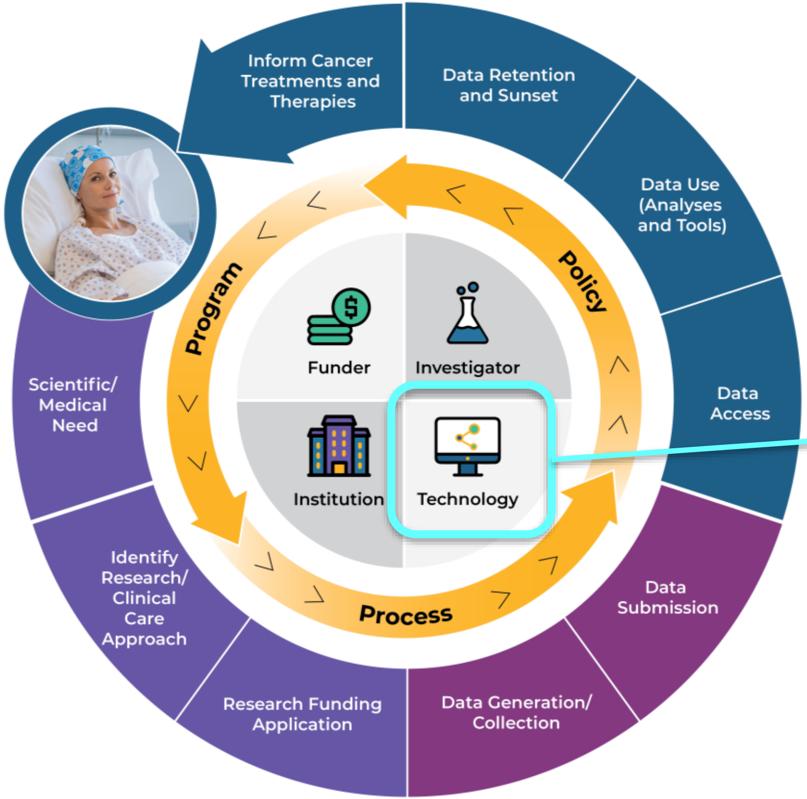
Policies to Promote Broad Use

Implementation of aggressive data management, sharing and access policies that ensure rapid, free and immediate access to all types of data.

Infrastructure to Support FAIR Principles

Technology platforms and tools that employ standards to make data findable, accessible, interoperable and reusable.

Sharing Data Openly through a Cancer Data Ecosystem



Inform Cancer Treatments and Therapies

Data Retention and Sunset

Data Use (Analyses and Tools)

Data Access

Data Submission

Data Generation/ Collection

Research Funding Application

Identify Research/ Clinical Care Approach

Scientific/ Medical Need

Program

Policy

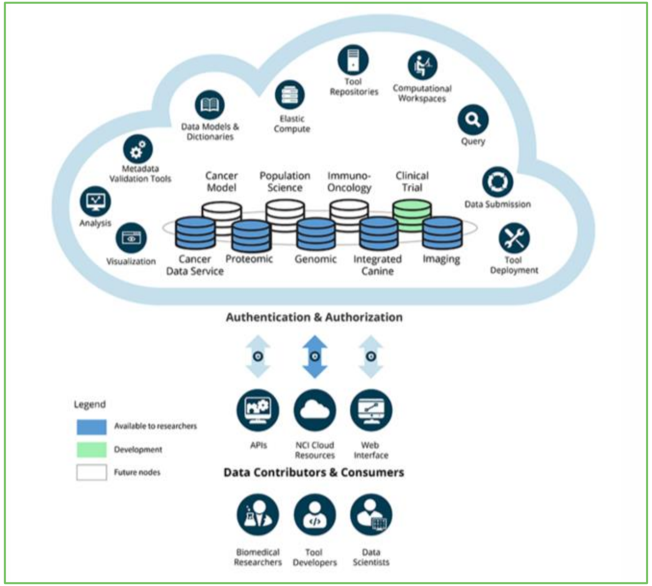
Process

Funder

Investigator

Institution

Technology



Tool Repositories

Computational Workspaces

Elastic Compute

Data Models & Dictionaries

Query

Metadata Validation Tools

Cancer Model

Population Science

Immuno-Oncology

Clinical Trial

Data Submission

Analysis

Visualization

Cancer Data Service

Proteomic

Genomic

Integrated Canine

Imaging

Tool Deployment

Authentication & Authorization

Legend

🟦 Available to researchers

🟩 Development

□ Future nodes

APIs

NCI Cloud Resources

Web Interface

Data Contributors & Consumers

Biomedical Researchers

Tool Developers

Data Scientists

Cancer Research Data Ecosystem

National Data Ecosystem: Integrating Cancer Research

SEER

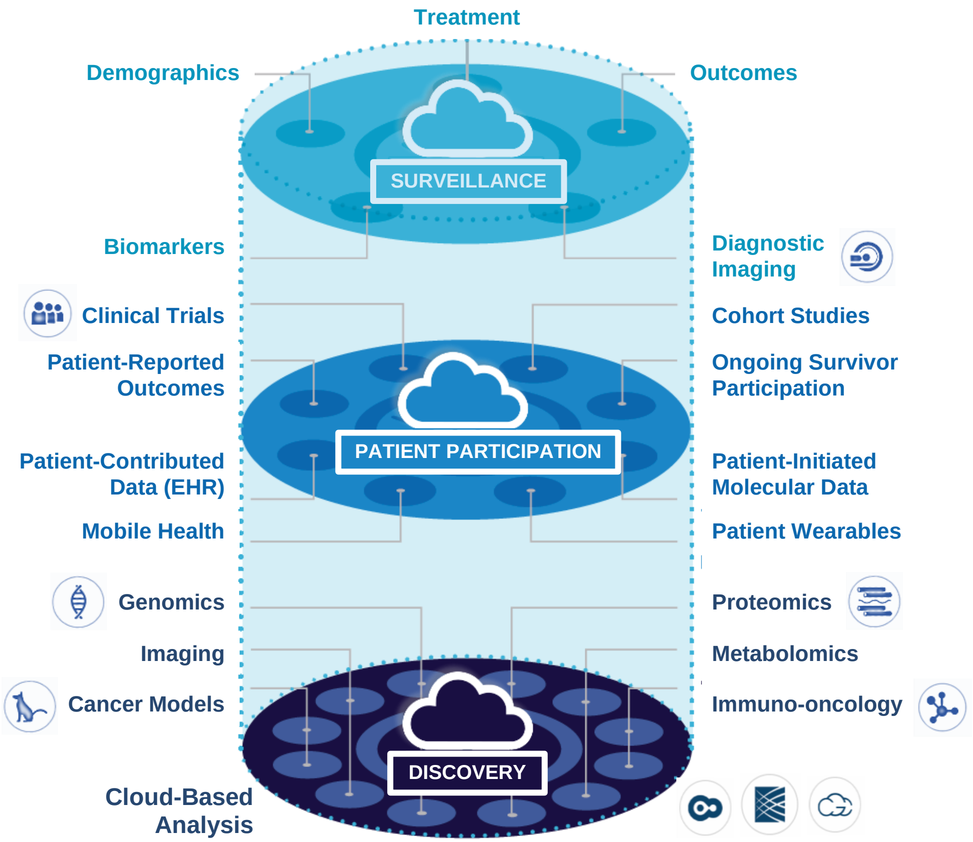
AACR American Association for Cancer Research FINDING CURES TOGETHER

PROJECT GENIE Genomics Evidence Neoplasia Information Exchange

APOLLO

HTAN HUMAN TUMOR ATLAS NETWORK

THE CANCER GENOME ATLAS



Treatment

Demographics

Outcomes

SURVEILLANCE

Biomarkers

Diagnostic Imaging

Clinical Trials

Cohort Studies

Patient-Reported Outcomes

Ongoing Survivor Participation

PATIENT PARTICIPATION

Patient-Contributed Data (EHR)

Patient-Initiated Molecular Data

Mobile Health

Patient Wearables

Genomics

Proteomics

Imaging

Metabolomics

Cancer Models

Immuno-oncology

DISCOVERY

Cloud-Based Analysis

CCDI BUILDING A COMMUNITY CENTERED AROUND CHILDHOOD CANCER CARE AND RESEARCH DATA

DATA TYPES: CLINICAL TREATMENT OUTCOME MOLECULAR BIOSPECIMEN LONGITUDINAL POPULATION

LEARN FROM EVERY CHILD

Improving the quality, consistency, and accessibility of data to make it easier for researchers to develop new and better treatments for children with cancer.

NCI National Clinical Trials Network a National Cancer Institute program

CPTAC National Cancer Institute at the National Institutes of Health

NCI National Cancer Institute

DOE Department of Energy

The Cancer Research Data Commons (CRDC)

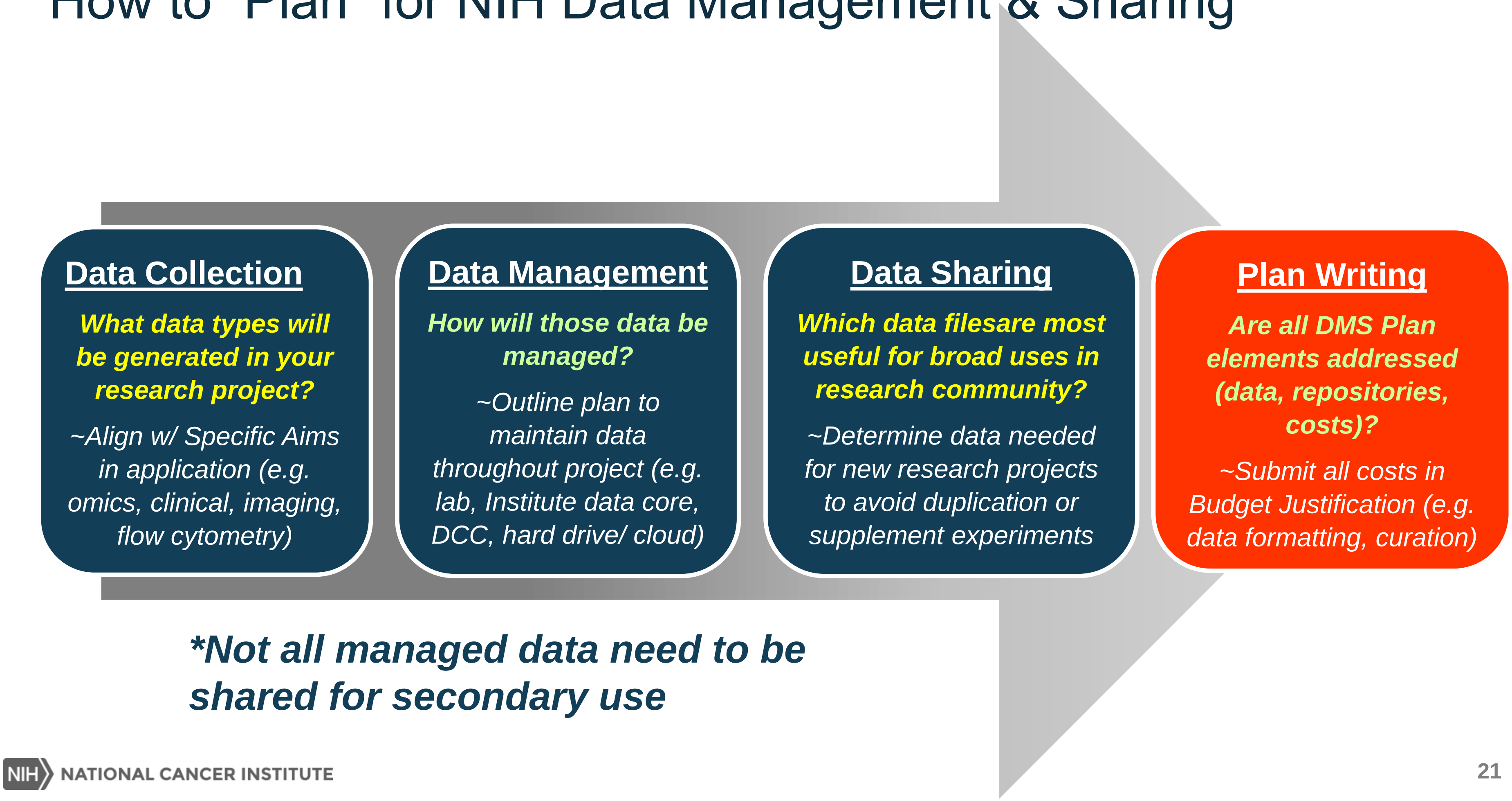
NIH Guidance for the New DMS Policy

Recommended DMS Plan Elements

Selecting Data Repositories

Allowable Costs of Data Management & Sharing

How to “Plan” for NIH Data Management & Sharing



Data Collection

What data types will be generated in your research project?

~Align w/ Specific Aims in application (e.g. omics, clinical, imaging, flow cytometry)

Data Management

How will those data be managed?

~Outline plan to maintain data throughout project (e.g. lab, Institute data core, DCC, hard drive/ cloud)

Data Sharing

Which data filesare most useful for broad uses in research community?

~Determine data needed for new research projects to avoid duplication or supplement experiments

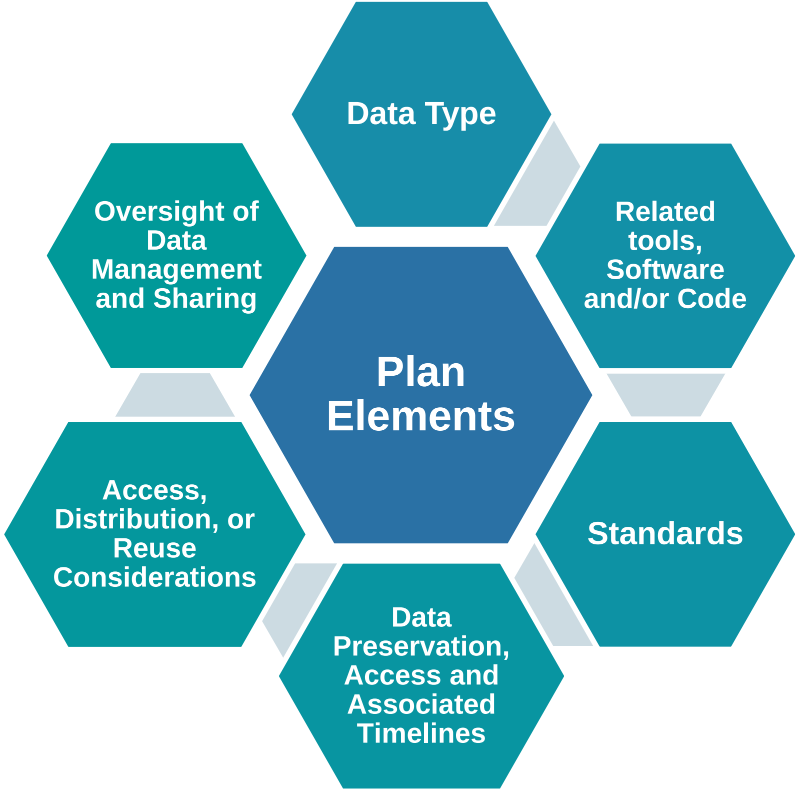
Plan Writing

Are all DMS Plan elements addressed (data, repositories, costs)?

~Submit all costs in Budget Justification (e.g. data formatting, curation)

****Not all managed data need to be shared for secondary use***

Six Elements of A DMS Plan

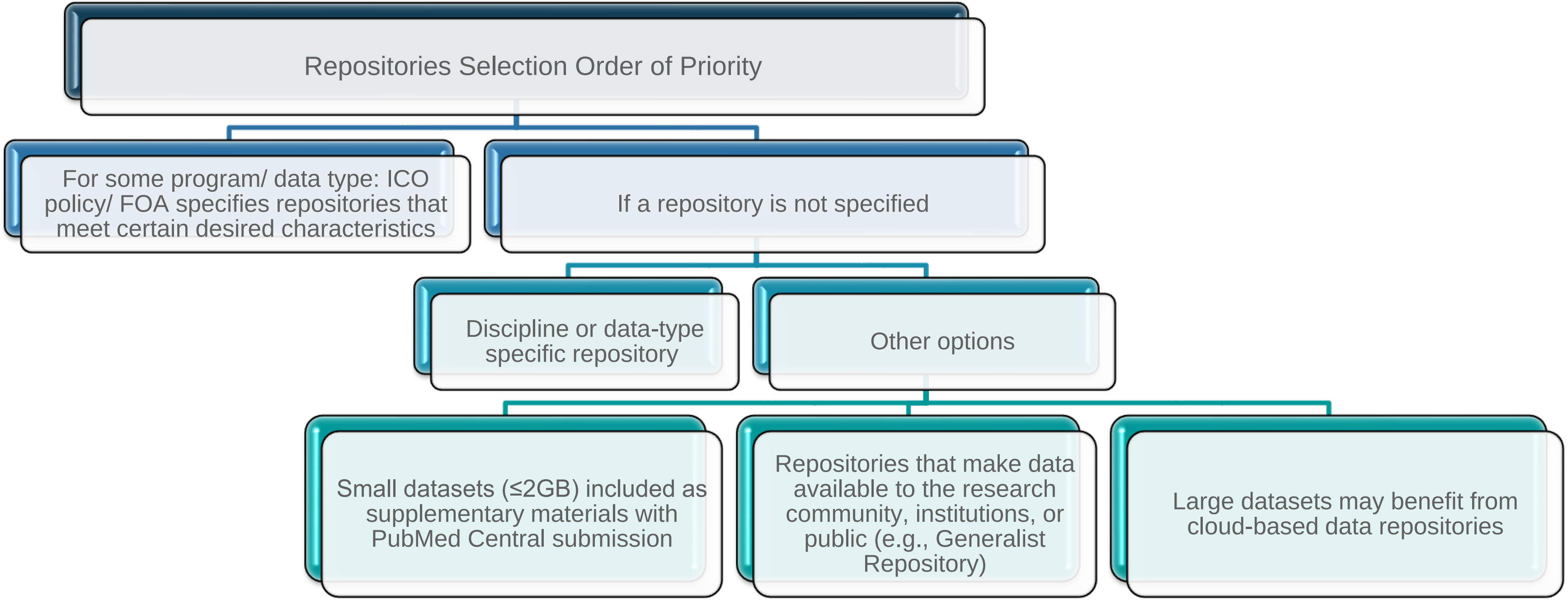

Data Type
Oversight of Data Management and Sharing
Related tools, Software and/or Code
Plan Elements
Access, Distribution, or Reuse Considerations
Standards
Data Preservation, Access and Associated Timelines

See Writing a Data Management & Sharing Plan for details

- **Data type**
  - Identifying data to be preserved and shared and listing metadata
- **Standards**
  - Standards to be applied to scientific data and metadata (i.e., data formats, data dictionaries, data identifiers, definitions, unique identifiers, and other data documentation)
- **Data preservation, access, timelines**
  - Repository to be used, persistent unique identifier, and when/ how long data will be available
- **Access, distribution, reuse considerations**
  - Description of factors for data access, distribution, or reuse
- **Related tools, software, code**
  - Tools and software needed to access and manipulate data
- **Oversight of data management and sharing**
  - Indicates how compliance with the DMS plan will be monitored and managed

Selecting A Data Repository

Promoting the use of established data repositories to improve FAIRness of data



Repositories Selection Order of Priority

For some program/ data type: ICO policy/ FOA specifies repositories that meet certain desired characteristics

If a repository is not specified

Discipline or data-type specific repository

Other options

Small datasets (≤2GB) included as supplementary materials with PubMed Central submission

Repositories that make data available to the research community, institutions, or public (e.g., Generalist Repository)

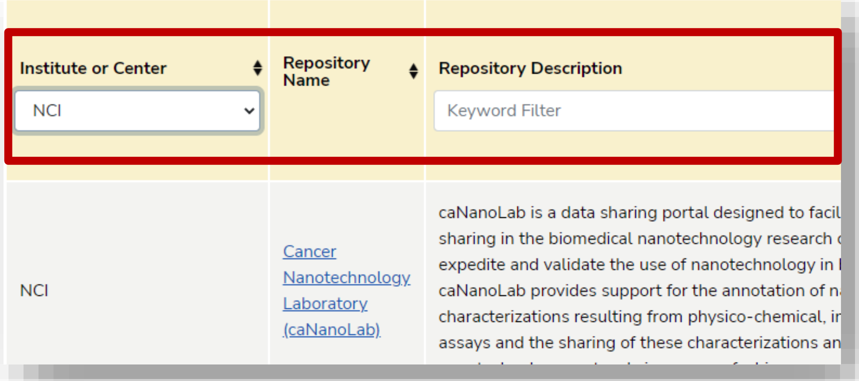
Large datasets may benefit from cloud-based data repositories

What is not a Repository? Supplemental material in a journal or Lab website

NIH Resources to Guide Repository Selection

NIH-Supported Repositories

- Filterable list of 75+ NIH Repositories that are open to taking in data of various types & formats
- PubMed holds up to 2Gb of data (as supplemental information to manuscripts)



Institute or Center	Repository Name	Repository Description
NCI		Keyword Filter
NCI	Cancer Nanotechnology Laboratory (caNanoLab)	caNanoLab is a data sharing portal designed to facil sharing in the biomedical nanotechnology research expedite and validate the use of nanotechnology in caNanoLab provides support for the annotation of n characterizations resulting from physico-chemical, ir assays and the sharing of these characterizations an

Other Repository Resources

- Generalist repositories (9 listed by NIH as part of GREI)
- Nature's Data Repository Guidance
- Registry of Research Data Repositories

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

See Repositories for Sharing Scientific Data

Data Management and Sharing Costs

ALLOWABLE COSTS

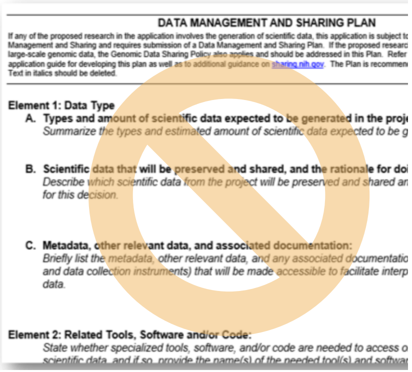
- Curating data/developing supporting documentation
- Preserving/sharing data through repositories
- Local data management considerations
- ***IMPORTANT*****:** Must be incurred during the performance period

UNALLOWABLE COSTS

- Infrastructure costs typically included in indirect costs
- Costs associated with the routine conduct of research (e.g., costs of gaining access to research data)
- ***Additional Example*****:** Data storage costs in NIH/ NCI-supported data repositories

See Budgeting for Data Management & Sharing for details

Responsibilities and Expectations: Peer Reviewers



DATA MANAGEMENT AND SHARING PLAN

If any of the proposed research in the application involves the generation of scientific data, this application is subject to
Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed researc
large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer
application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommen
Text in italics should be deleted.

Element 1: Data Type

A. **Types and amount of scientific data expected to be generated in the proje**
Summarize the types and estimated amount of scientific data expected to be g

B. **Scientific data that will be preserved and shared, and the rationale for doi**
Describe which scientific data from the project will be preserved and shared an for this decision.

C. **Metadata, other relevant data, and associated documentation:**
Briefly list the metadata, other relevant data, and any associated documentatio and data collection instruments) that will be made accessible to facilitate interp data.

Element 2: Related Tools, Software and/or Code:
State whether specialized tools, software, and/or code are needed to access o scientific data, and if so, provide the name(s) of the needed tool(s) and softwar

- Peer Review Will NOT assess DMS Plans*
  - DMS Plan assessment is an administrative issue, reviewers assess only scientific merit of applications
- Peer Review Will NOT see full Plan or be providing comments on Plans*
  - DMS Plan will be excluded from the application image shown to peer reviewers
- Peer Review will NOT factor DMS Plans into their impact scores*

**Exceptions for FOAs where data sharing is core to the science, e.g., FOAs for data dissemination centers)*

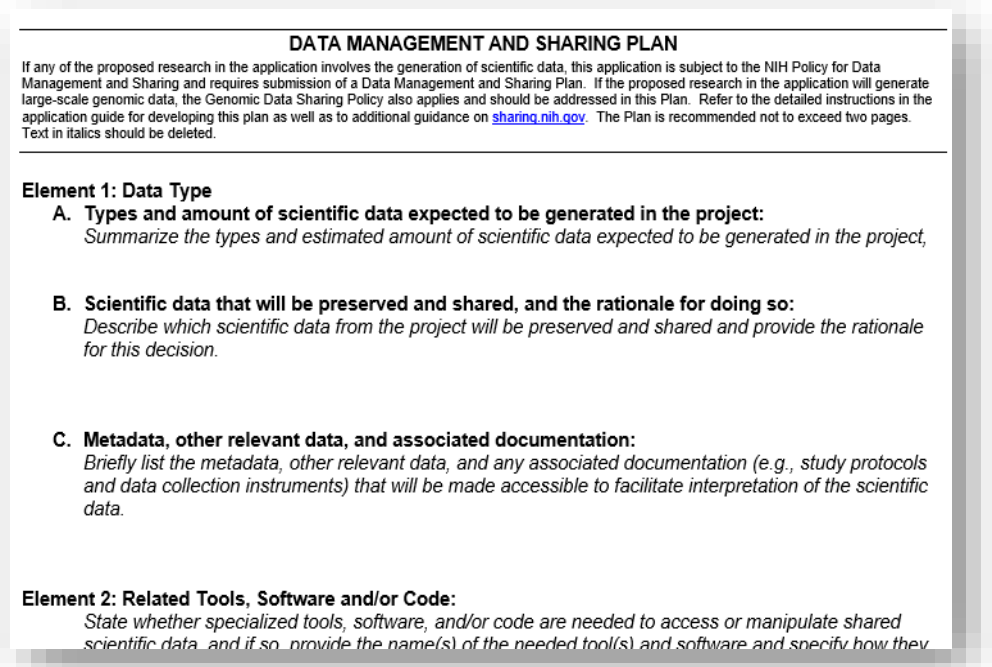
What Should a DMS Plan Look Like (OER)?

✓ Plans should be no more than 2 pages in length

✓ Applications subject to both DMS and GDS Policies submit a single Plan

✓ Optional format page available

✓ Exploring structured form in future *(NCI & NICHD have each developed forms; possible pilot)*



DATA MANAGEMENT AND SHARING PLAN

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommended not to exceed two pages. Text in italics should be deleted.

Element 1: Data Type

A. **Types and amount of scientific data expected to be generated in the project:**
Summarize the types and estimated amount of scientific data expected to be generated in the project.

B. **Scientific data that will be preserved and shared, and the rationale for doing so:**
Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.

C. **Metadata, other relevant data, and associated documentation:**
Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

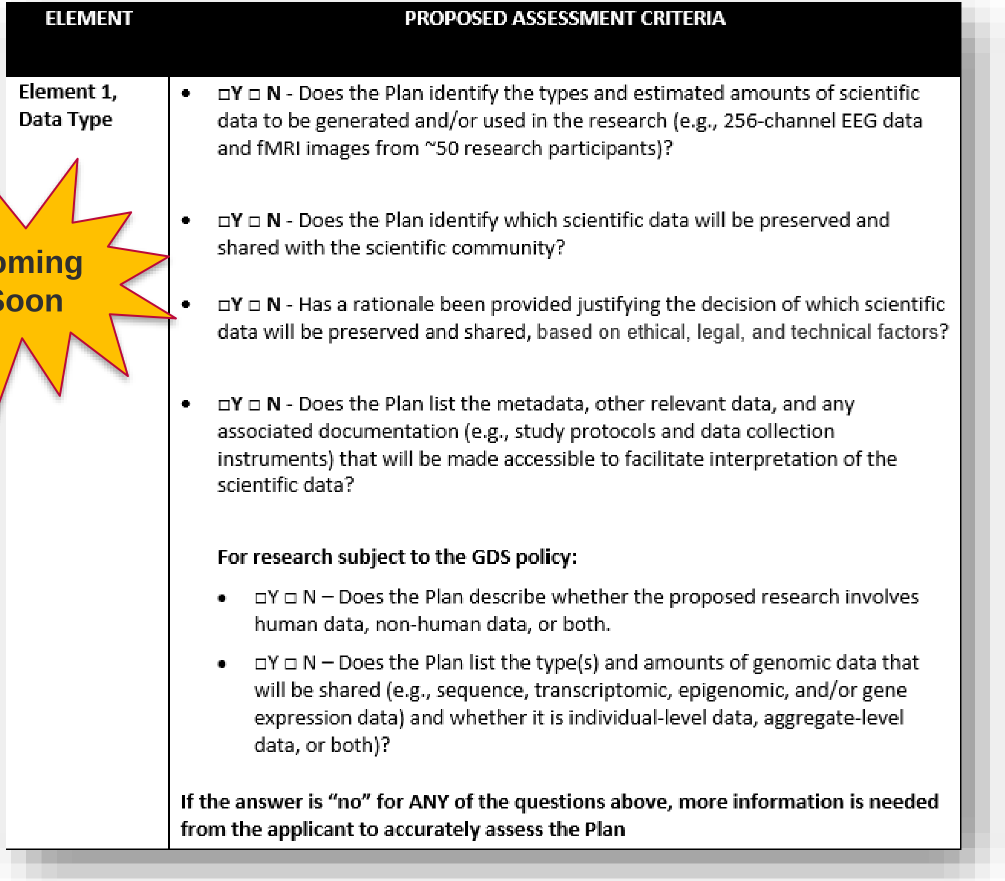
Element 2: Related Tools, Software and/or Code:
State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they

DMS Plan format page will be added to list of Format Pages and incorporated into FORMS-H application instructions by Fall 2022

A Decision Support Tool to Guide Plan Assessment

- Optional tool to guide POs through DMS Plan assessment
- Stepwise list of considerations for each element
- Additional ICO or program-specific guidance may also apply

Coming Soon



ELEMENT	PROPOSED ASSESSMENT CRITERIA
Element 1, Data Type	• □**Y** □ **N** - Does the Plan identify the types and estimated amounts of scientific data to be generated and/or used in the research (e.g., 256-channel EEG data and fMRI images from ~50 research participants)? • □**Y** □ **N** - Does the Plan identify which scientific data will be preserved and shared with the scientific community? • □**Y** □ **N** - Has a rationale been provided justifying the decision of which scientific data will be preserved and shared, based on ethical, legal, and technical factors? • □**Y** □ **N** - Does the Plan list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data? **For research subject to the GDS policy:** • □Y □ N – Does the Plan describe whether the proposed research involves human data, non-human data, or both. • □Y □ N – Does the Plan list the type(s) and amounts of genomic data that will be shared (e.g., sequence, transcriptomic, epigenomic, and/or gene expression data) and whether it is individual-level data, aggregate-level data, or both)? **If the answer is “no” for ANY of the questions above, more information is needed from the applicant to accurately assess the Plan**

Harmonization of DMS and Genomic Data Sharing Plans

Investigators	**Program Staff**
• One Plan: Genomic data sharing considerations to be addressed in DMS Plans using the expected DMS Plan elements. • Timeline: At the time of application or Just-in-Time. • *For human data subject to GDS*: Applicants should complete the DMS Plan anticipating sharing according to the criteria in the Institutional Certification (IC).	▪ IC or provisional certification must be submitted and accepted before issuing the award. Investigators should state in the DMS Plan what data can be shared and how if IC criteria cannot be met. ▪ Plans will be reviewed by Program Staff. Peer reviewers will no longer comment on the Plans but rather on the reasonableness of the DMS budget. ▪ Submission timelines for non-human and human data subject to GDS will remain unchanged. ▪ Compliance and enforcement terms for awards subject to GDS will be handled in accordance with that under DMS (i.e., “end of the performance period” is the *latest* opportunity to submit data).

Notice Number: NOT-OD-22-198

How Do the DMS and Genomics Data Policies Compare?

<table>
<tr><th></th><th>2015 NIH GDS Policy</th><th>2023 NIH DMS (New Policy)</th></tr>
<tr><td><b>Effective Dates</b></td><td>▪ <b>New applications</b>: received on/after January 25, 2015</td><td>▪ <b>New applications & Competitive renewals</b> (Type 2, previously funded awards): For submission on/after January 25, 2023</td></tr>
<tr><td rowspan="2"><b>Scope</b></td><td>▪ <b>Large-scale genomics data</b> &amp; some smaller studies <i>(programmatic priority, rare diseases)</i>
▪ <b>Thresholds</b> &amp; <b>assay types</b> defined</td><td>▪ All NIH-supported research that generates <b>scientific data</b>
▪ Not applicable to other activities (e.g. training, infrastructure development)
▪ No detailed expectations defined outside of GDS, CT or programs</td></tr>
<tr><td colspan="2">▪ Intramural & Extramural research mechanism (grants, cooperative agreements, contracts, other transactions)
▪ Human and non-human data; no budget threshold</td></tr>
<tr><td rowspan="2"><b>NIH Expects <i>(for PIs)</i></b></td><td>▪ Submit & release data through NIH repositories
▪ Use guidelines for standard formats, levels of data</td><td>▪ <b>Prospectively plan</b> how to preserve & share scientific data
▪ Outline in <b>Data Management & Sharing (DMS) Plan</b> (w/ applications), and
▪ Report data sharing progress in annual submission of <b>RPPR</b></td></tr>
<tr><td colspan="2"><b><i>*Only one plan will be submitted</i>; DMS plans</b> will <b>include genomics</b> elements along with other data types (i.e. no separate GDS plan submitted after January 25, 2023)</td></tr>
<tr><td><b>Timelines</b></td><td>▪ Lower-level primary data: released by <b>9 months</b> (full, QC dataset generation)
▪ Summary analyses at publication</td><td>▪ Shared scientific data: <b>as soon as possible</b> → by time of associated publication, or end of performance period (whichever comes first)</td></tr>
</table>