DOE Public Access and Data Management

ASCAC Meeting
January 2025

Michael Cooke, PhD
Senior Technical Advisor
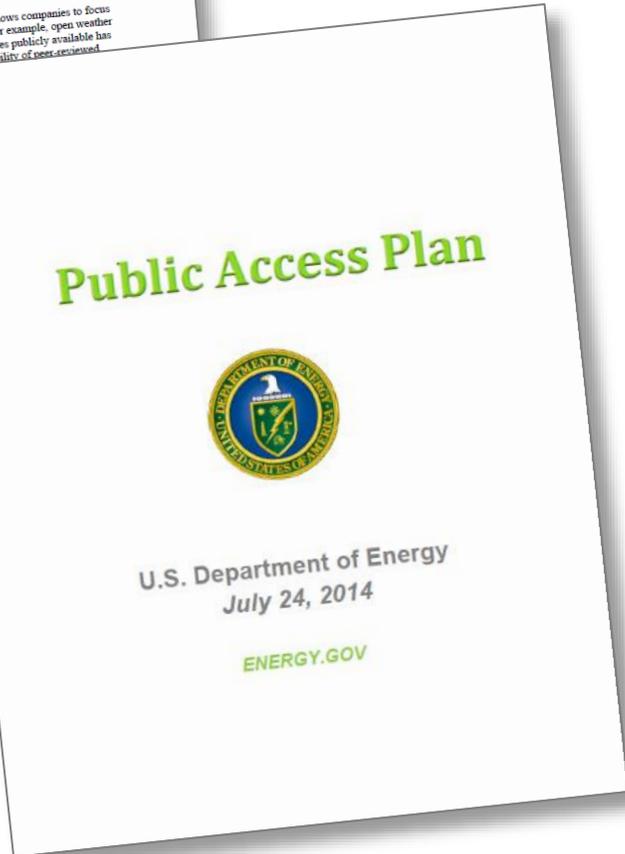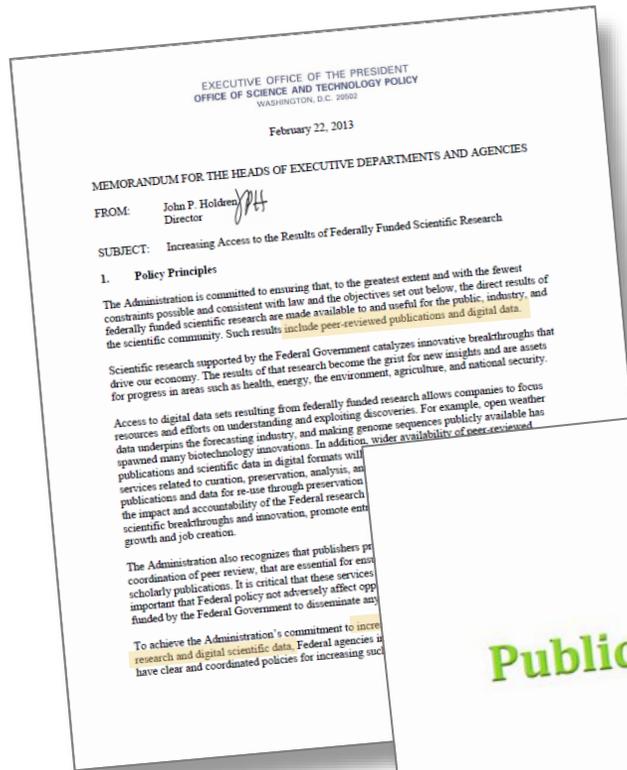Office of the Deputy Director for Science Programs

U.S. Department of ENERGY | Office of Science

Energy.gov/science

2013 – 2014 Public Access

- 2013 Public Access Memo issued by the White House Office of Science and Technology Policy (OSTP)
  - Results of federally funded scientific research, including peer-reviewed publications and digital research data, should be made publicly available
  - Allowed 1-year embargo of peer-reviewed articles
- 2014 DOE Public Access Plan
  - Requires author submission of accepted manuscript for peer-reviewed publications to DOE within 12 months of publication
  - Government purpose license used to share manuscripts through DOE PAGES® with voluntary participation of publishers
  - Data Management Plan (DMP) requirements for public sharing of digital research data

August 2022 Nelson Memo

All federal science agencies, including DOE, required to develop new Public Access Plans

- Significant updates include:
  - Emphasis on use/re-use; machine readability; equitable access
  - Immediate public access to publications, removing 12-month embargo
  - Immediate public access to data displayed in or underlying publications
  - Expanded use of persistent identifiers (PIDs)
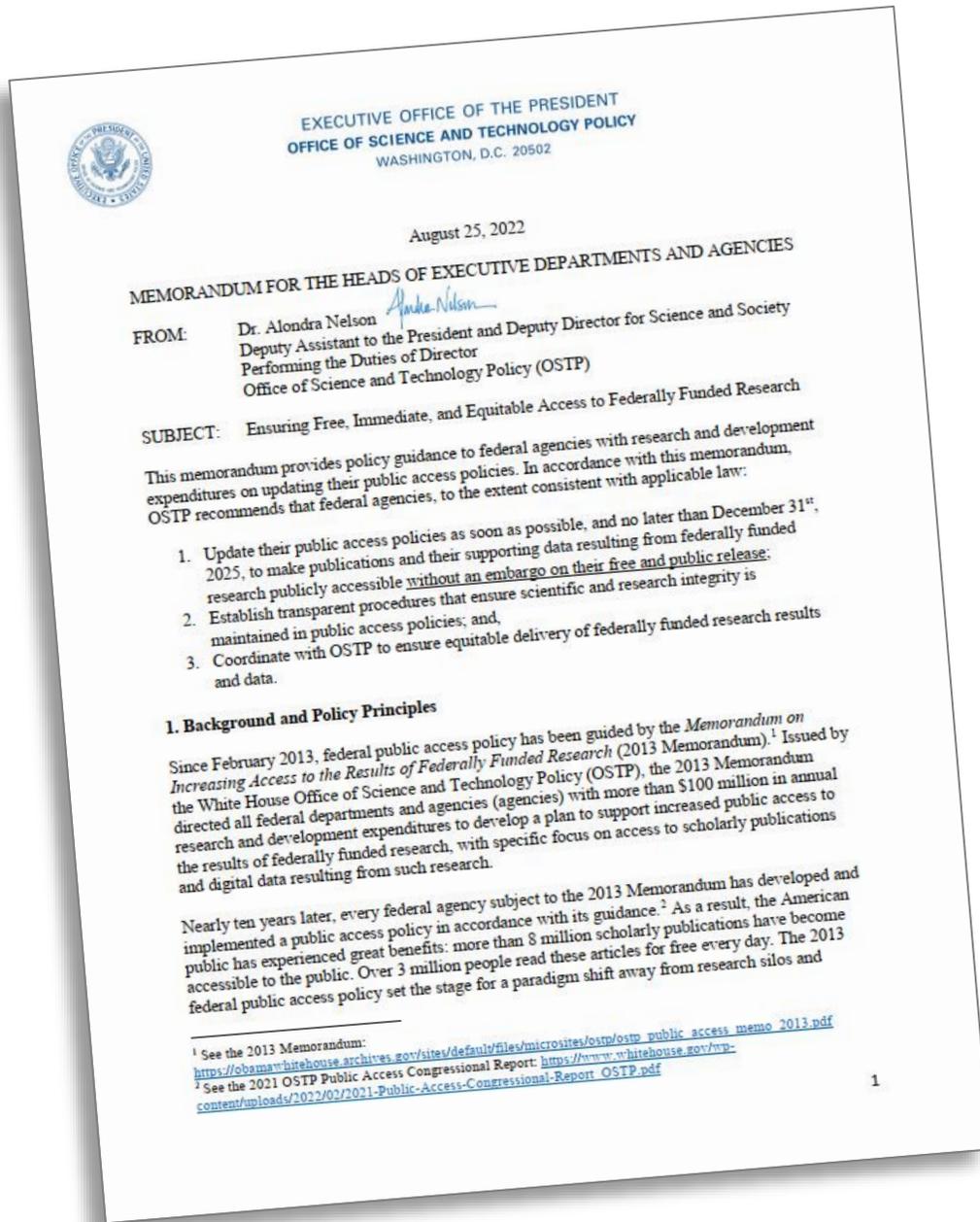
Implementation Timeline

2022 OSTP Public Access Memo Section Descriptions

Section 3: Publications & Data   Section 4: PIDs to Ensure Research & Scientific Integrity   Section 5: Interagency Coordination

Aug 25, 2022	Feb 21, 2023	Dec 31, 2024	Dec 31, 2025	Dec 31, 2026	Dec 31, 2027

OSTP Public Access Policy Guidance released

Section 3: DOE Public Access Plan due to OSTP/OMB

Section 3: Last date to publish related DOE policies

Section 4: Last date to provide optional DOE Public Access Plan update to OSTP/OMB

Section 3: Last date for related policies to be effective

Section 4: Last date to publish related DOE policies

Section 4: Last date for related policies to be effective

Section 5: Ongoing interagency coordination

Development of DOE's New Public Access Plan – Released June 2023
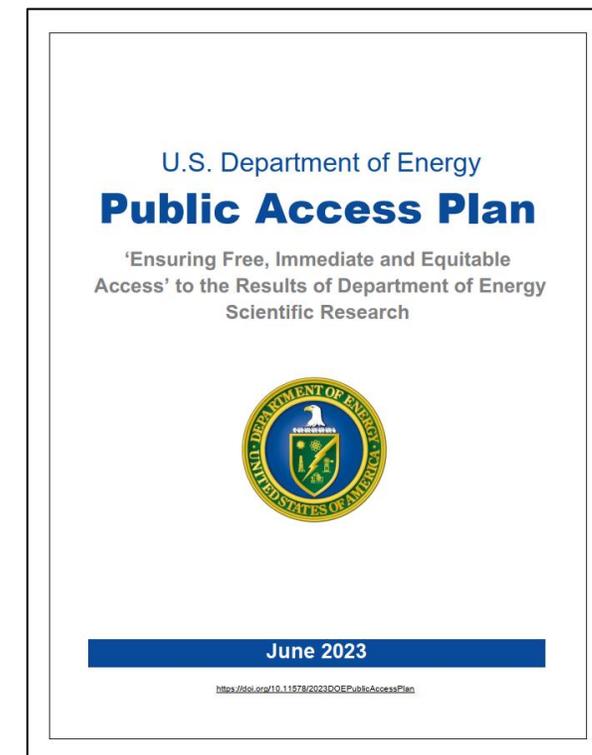
Intra-Agency Coordination
- DOE-wide participation (EERE, FE, NE, OE, ARPA-E, MA, GC, NNSA)
- SC-led author team
- Coordinated with DOE and SC Working Groups on Digital Data
- DOE researcher community input through Labs' STI managers

Interagency Coordination
- OSTP Subcommittee on Open Science (SOS); SC co-chairs three SOS working groups
- Persistent Identifier Services partners from 12 agencies

External Community Engagement
- Professional societies
- Publishers
- Libraries

U.S. Department of Energy
Public Access Plan
'Ensuring Free, Immediate and Equitable Access' to the Results of Department of Energy Scientific Research

June 2023

https://doi.org/10.11578/2023DOEPublicAccessPlan

Full plan and FAQ available at:
https://www.energy.gov/doe-public-access-plan
https://doi.org/10.11578/2023DOEPublicAccessPlan

2023 DOE Public Access Plan

Publications

- Move from 12-month embargo to immediate access upon publication

- Continue to submit accepted manuscripts via E-Link, but earlier in reporting process

- Provide access through DOE's designated repository, DOE PAGES®

- Emphasize author deposits of accepted manuscripts (green OA) – DOE

Data

- Now Data Management and Sharing Plans (DMSPs)

- "Scientific Data" to validate and replicate research findings

- Data underlying publications should be made available at time of publication

- Timeline for sharing other scientific data

- Repository selection should align with NSTC Desirable Characteristics of Data Repositories guidance
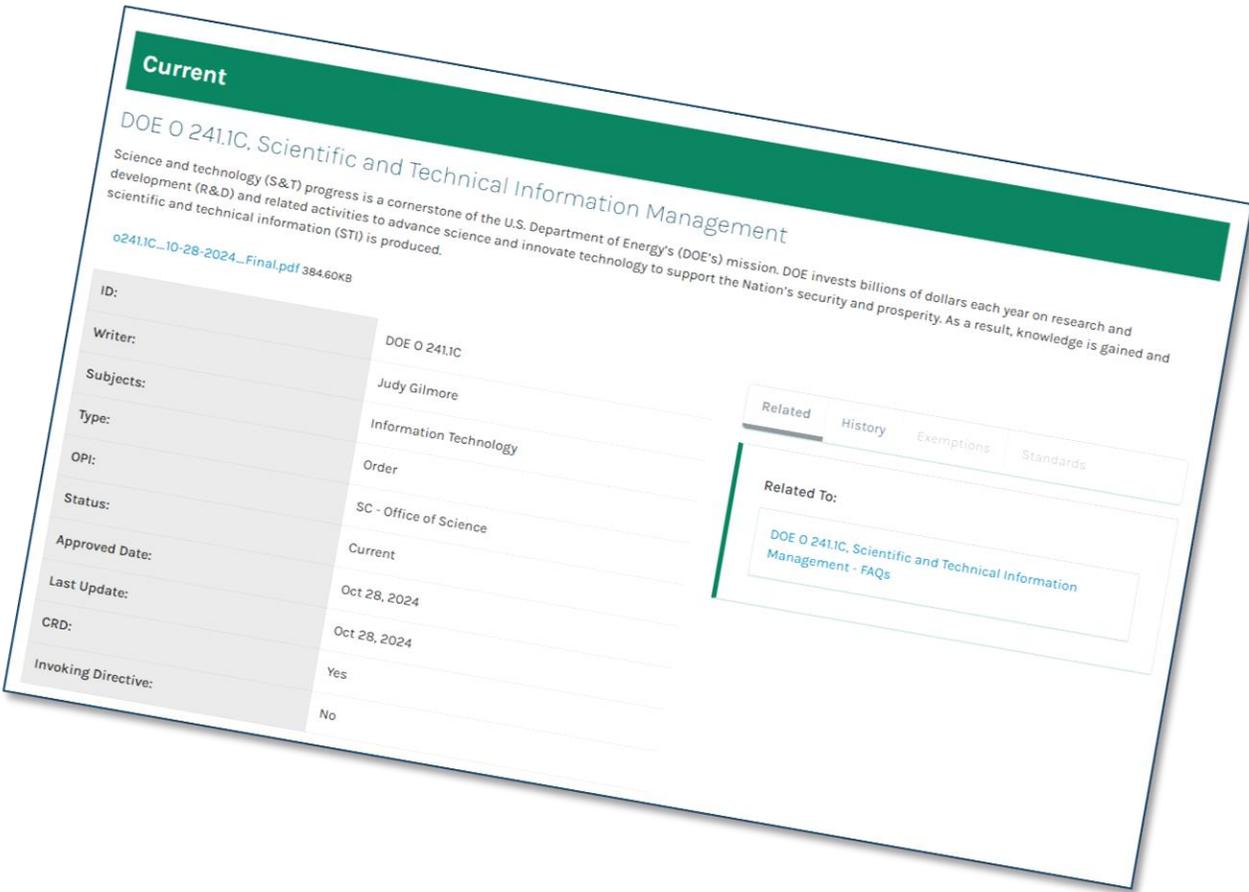
Persistent Identifiers

- Collect metadata associated with publications and data

- Metadata to include authors, affiliations and funding with associated PIDs, publication date, and PID for output

- Instruct researchers to obtain a PID for themselves and use when publishing and reporting R&D outputs

- Researcher PIDs must meet common/core standards

- PIDs for awards

2023 DOE Public Access Plan: https://www.energy.gov/doe-public-access-plan

U.S. Department of **ENERGY** | Office of Science

Energy.gov/science

DOE Order 241.1C: Scientific and Technical Information Management

- Order ensures DOE fulfills its scientific and technical information (STI) management responsibilities
  - Requires identification, collection, review, control, preservation, and dissemination of R&D outputs/STI
  - Establishes requirements and roles for federal staff and contractors

- Order updated in October 2024:
  - Policy mechanism used to implement DOE Public Access Plan – addresses publication, data management, and PID requirements
  - Order also addresses S&T Risk Matrix, National Security Presidential Memo 33 PID requirements, and DOE O 471.7 Controlled Unclassified Information
  - Series of rollout webinars underway to engage with laboratories and DOE programs

DOE Order 241.1C: https://www.directives.doe.gov/directives-documents/200-series/0241.1-border-c

U.S. Department of ENERGY | Office of Science

Energy.gov/science

Public Access to Publications Implementation

Implementation Mechanisms	E∞LINK	DEPARTMENT OF ENERGY DOE PAGES® Public Access Gateway for Energy & Science
Published revision to DOE O 241.1C and Contractor Requirements Document (CRD)	Accepted Manuscripts continue to be submitted in E-Link (expected earlier), including full text pdf	Public Access provided through DOE's designated repository, DOE PAGES®
Updating NOFO, award term, financial assistance checklist, and FA guide	Requirement in effect to submit journal article accepted manuscripts to OSTI so they can be made available without delay after publication	Removal of 12-month embargo of full text
Socialize requirements – webinars, trainings	Requirement in effect for labs and financial assistance, but in a transition period to adjust to new requirement and workflow	

Pathways to Public Access

DOE allows for
<u>author's choice</u> in publishing

Green Open Access (OA) is encouraged
but Gold OA is allowed if fees are
"reasonable"

U.S. Department of **ENERGY** | Office of Science

Energy.gov/science

Current DOE Data Management Overview

DOE data management principles		
Enable discovery	Share, preserve, validate	Cost management

DOE Data Management Plan (DMP) requirements			
Share, preserve, validate	Make data associated with publications accessible	Availability of data management resources	Privacy, security, confidentiality

- DMPs are reviewed, but there is flexibility in the process used for collection and review
- Additional requirements may be identified in a solicitation or invitation for research funding
- Implementation is supported through commensurate budget for the approved DMP scope
- Current DMP requirements and guidance applicable through September 30, 2025

Current DOE guidance: *https://www.energy.gov/datamanagement/doe-requirements-and-guidance-digital-research-data-management*
Current program-specific guidance: *Office of Energy Efficiency and Renewable Energy, Office of Science, SBIR/STTR*

Data Management Requirements Updates

- Highlighted changes from current data management strategy to FY 2026
  - "Data Management Plans" will become "Data Management and Sharing Plans"
  - Updated principles emphasize aim to "maximize appropriate data sharing"
  - Reporting requirements aim to enhance compliance monitoring and evaluation metrics

Data Management Plan (current)	Data Management and Sharing Plan (FY 26)
• "Research Data" to validate research findings	• "Scientific Data" to validate and replicate research findings
• Data underlying publications should be made as accessible as possible	• Data underlying publications should be made available at time of publication
• No explicit requirement regarding other research data	• Timeline for sharing other scientific data
• No explicit requirement regarding repository selection	• Repository selection should align with NSTC Desirable Characteristics guidance *(next slide)*

Desirable Characteristics of Data Repositories

- Guidance by the National Science and Technology Council (NSTC) Subcommittee on Open Science for federally funded research
  - Improves consistency in instructions to researchers about selecting data repositories
  - Helps ensure research data are findable, accessible, interoperable, and reusable (FAIR) to the greatest extent possible, while integrating privacy, security, and other protections

Organizational Infrastructure	Digital Object Management	Technology	Additional Considerations for Human Data
• Free and Easy Access • Clear Use Guidance • Risk Management • Retention Policy • Long-term Organizational Sustainability	• Unique Persistent Identifiers • Metadata • Curation and Quality Assurance • Broad and Measured Reuse • Common Format • Provenance	• Authentication • Long-term Technical Sustainability • Security and Integrity	• Fidelity to Consent • Security • Limited Use Compliant • Download Control • Request Review • Plan for Breach • Accountability

Desirable Characteristics of Data Repositories for Federally Funded Research, guidance by the NSTC Subcommittee on Open Science, published May 2022

Data Management Implementation

2023 Public Access Plan: Scientific Data Management Principles		
Increase pace of scientific discovery	Protect integrity, enhance value of science	Maximize appropriate data sharing

2023 Public Access Plan: Data Management and Sharing Plan (DMSP) Requirements				
Validation and replication of results	Timely and equitable access	Data repository selection	Data management and sharing resources	Data sharing limitations

Implementation:
- All DOE-funded R&D awards and contracts will be subject to a DOE approved DMSP and require reporting of shared data
- DMSP implementation will be supported through commensurate budget for approved scope

- Data management planning and submitting data as STI requirements included in DOE O 241.1C update, Attachment 6
- Developing updates to award term, FARC, RPPR
- New DMSP requirements applicable for research projects submitted on or after October 1, 2025

Persistent Identifier (PID) Overview

From the White House Memos – *"A digital identifier that is globally unique, persistent, machine resolvable and processable, and has an associated metadata schema."*

A long-lasting, managed, and registered unique digital reference (often in the form of a URL) to a research object (e.g. person, organization, research output, award) that can be represented or described online.

PID Benefits
- Burden Reduction
- Disambiguation
- Proper Credit
- Discovery and Access
- Interoperability
- Analytics and Impact

PIDs for Research Outputs

(Publications, Reports, Data, Software)

https://doi.org/10.1016/j.rinp.2023.106511
https://doi.org/10.11578/dc.20230501.1

PIDs for People

(PIs, Researchers, Senior/Key Personnel)

https://orcid.org/0000-0002-8523-1478

PIDs for Organizations

(Funders, Universities, Publishers, Facilities)

https://ror.org/01bj3aw27

PIDs for Awards

(Grants, Contracts, Facility Use)

https://doi.org/10.46936/10.25585/60000014

Persistent Identifier Implementation – OSTI

PIDs for Outputs

- OSTI collects and assigns PIDs for R&D outputs – use DOIs.
- Collect DOIs for journal article accepted manuscripts assigned by publishers.
- OSTI provides DOI assignment services that are part of R&D output submission processes in E-Link.
- Assign DOIs to technical reports, data, and software if not assigned by another source.

PIDs for People

- OSTI has optionally collected ORCID iDs associated with R&D output authors.
- Continue to collect author ORCID iDs.
- Will be developing additional ORCID iD support in E-Link.
- Explore collecting author ORCID iD from other sources (in addition to being provided by PI/submitter).

PIDs for Organizations

- OSTI has an organization authority to standardize organization names and information.
- Developed new organization authority with organization PIDs.
- Organization PIDs will be added to STI metadata.
- Organizations will be able to submit organization PIDs with metadata associated with R&D outputs.

Persistent Identifiers Implementation

Next Steps

More details in backup...

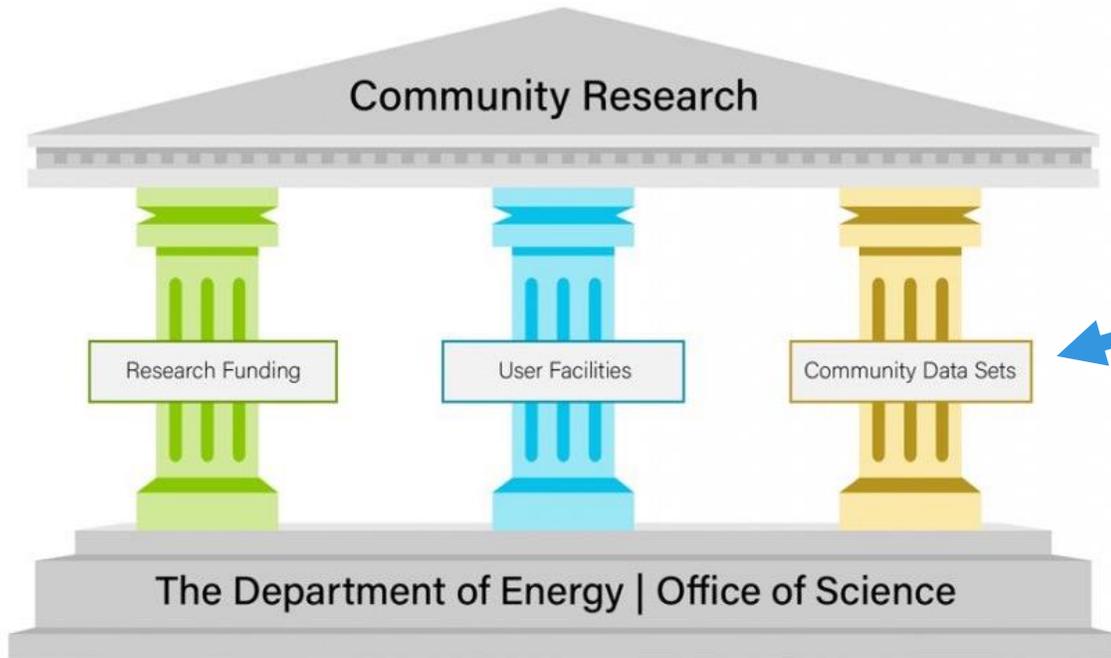Socialize PID requirements (DOE O 241.1C and FAL)

Ensure requirements are reflected in financial assistance guide to be included NOFOs and Award Terms

Update OSTI infrastructure to better support PIDs

PIDs will be included in STI training

Community conversations about what PID support is needed

DOE PuRe Data Resources

Public Reusable Research (PuRe) Data Resources aim to make data publicly available in order to advance scientific or technical knowledge

https://science.osti.gov/Initiatives/PuRe-Data

Champions of Open Science

- White House OSTP Year of Open Science Recognition Challenge
  - Recognizing open science stories to benefit society
  - 5 [winners announced](#) in March 2024

Project [Jupyter](#): reproducible and collaborative computational science and education

Category: Technical Advancement to Enable Open Science

Federal Support: NSF and DOE

Project leads: Brian Granger, Jason Grout, Fernando Pérez, Ana Ruvalcaba, and Steven Silvester

"Because Jupyter notebooks help to make computational research reproducible and accessible, Jupyter has enabled millions of researchers and thousands of organizations to adopt open science practices in their research and education."

Used by:

Energy.gov/science

Office of Science Initiatives: Broadening Participation

RENEW
Reaching a New Energy Sciences Workforce

FAIR
Funding for Accelerated, Inclusive Research

PIER Plans
Promoting Inclusive and Equitable Research

Integrated Research Infrastructure & HPDF

2023 IRI Architecture Blueprint Activity identified three science patterns:

- *Time-sensitive patterns*
- *Data-integration-intensive patterns*
- *Long-term campaign patterns*

- DOE envisions a revolutionary ecosystem, the **Integrated Research Infrastructure**, to deliver seamless, secure interoperability across National Lab facilities

- The **High Performance Data Facility (HPDF)** will enable analysis, preservation, and accessibility to the staggering amounts of experimental data produced by SC facilities

  - Distributed operations model will be essential to long-term success and required performance levels
  - Hub & spoke architecture will provide seamless, tailored service to users worldwide
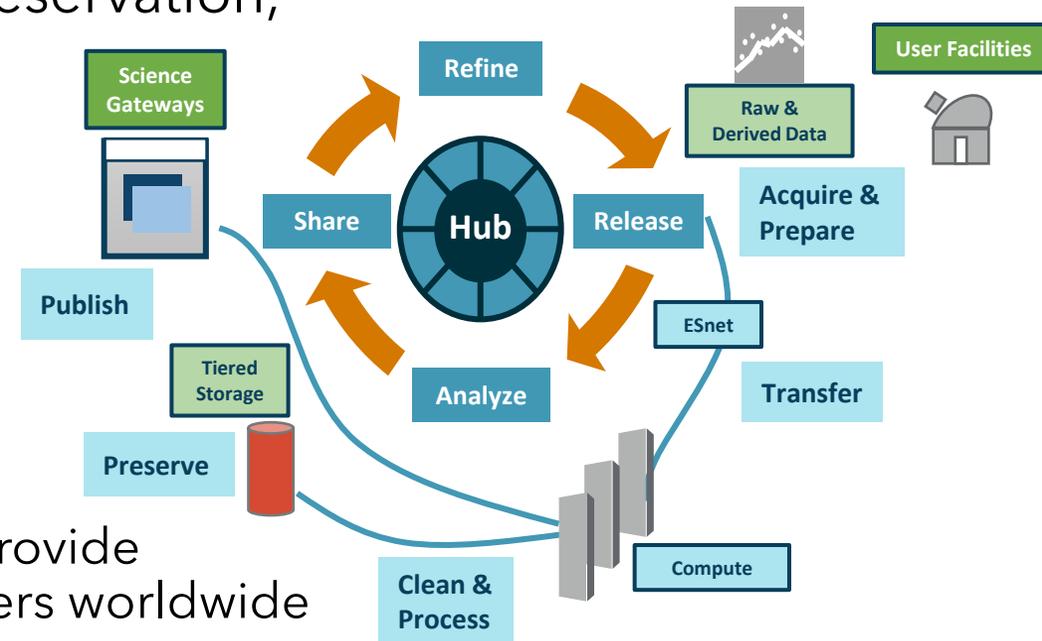
NSTC Report: Framework for Considering Data Infrastructure and Interconnectivity

- Developed by the NSTC Subcommittee on R&D Infrastructure's (RDI's) Data Infrastructure Working Group
- Provides a high-level framework for considering data infrastructure and interconnectivity during planning, developing, operating, upgrading, and assessing RDI
- Identifies several prevalent challenge areas often faced when developing and upgrading data infrastructure
- Proposes collective agency and community actions to disseminate practices and approaches for planning data infrastructure and addressing common challenge areas

Framework: Principal Considerations Areas for Data Infrastructure Design	Cross-cutting Challenge Areas	Recommendation Areas for Collective Action
• Science goals and mission priorities for data • Users and utilization • Data inventory, management, and stewardship • Dynamic data ecosystem • Project governance and partnering	• Transferring and managing large-scale data • Data integration and data infrastructure interoperability • Operating in the commercial cloud • Handling sensitive and secure data • Workforce development and nurturing	• Dissemination and exchange of practices • Coordination • Workforce

Report link: https://www.whitehouse.gov/ostp/news-updates/2024/12/17/nstc-framework-for-considering-data-infrastructure-and-interconnectivity/

Vision and Burning Questions

- Office of Science vision for implementing the DOE Public Access Plan
    - Enable researchers to develop & implement responsive DMSPs in partnership with data, computing, and networking experts
    - Integrate PIDs into the scientific workflow to connect research outputs, including publications, data, code, and other tools
    - Encourage data citation and reuse to enable replication of results and new modes of science
    - Enable career-impacting recognition for contributing public data and tools of open science

- Burning questions
    - How can we enable researchers to develop and implement responsive DMSPs while letting them focus on their science?
        - Through SC infrastructure? Through helpful guidance? Through community best practices?
    - How can we enable scientists to explore new opportunities made possible by FAIR data?
        - While also recognizing impactful data/code/tools and the people behind them?
    - How can we lower the barrier for broadening participation in science?

U.S. Department of ENERGY | Office of Science

E-Link

- DOE system for collecting R&D results/STI (including scholarly publications/journal article accepted manuscripts)
- Financial assistance recipient PI's use system to submit DOE-funded R&D outputs (requirements included in award term and detailed in financial assistance reporting checklist)
- Lab researchers submit R&D outputs through lab review and release systems – lab systems provide information to OSTI/E-Link
- E-Link is developed and managed by OSTI
- A new E-Link 2.0 version to better support public access and PIDs will be released in 2025

E-Link: https://www.osti.gov/elink/forms.jsp

Energy.gov/science

DOE PAGES

- Search tool for finding DOE-funded journal article publications and accepted manuscripts

- Publication records submitted in E-Link are discoverable in DOE PAGES

- Content indexing in common search engines to widely disseminate DOE-funded publications

- Historically, had 12-month delay before making the accepted manuscript full text publicly available

DOE PAGES: https://www.osti.gov/pages/

Energy.gov/science

Persistent Identifier Implementation – DOE O 241.1C

DOE federal and contractor employees conducting R&D work must obtain a persistent identifier (PID) for themselves that meets the common/core standards specified in the NSPM-33 Implementation Guidance or successor guidance (e.g., ORCID iD).

The PID must be used by these employees in published research outputs when available and be provided to OSTI with STI metadata records.

A PID (e.g., digital object identifier [DOI]) must be associated with the following STI records:

1) Accepted Manuscripts/Journal Articles – If a PID is assigned by the publisher, it must be provided to OSTI with the metadata record.

2) Scientific Data that are publicly accessible (i.e., posted on a public website or repository) – If the data has a PID, it must be provided to OSTI with the metadata record. If a PID is not provided, OSTI will assign one.

3) Technical Reports with no distribution limitations – If the report has a PID, it must be provided to OSTI with the metadata record. If a PID is not provided, OSTI will assign one.

4) Scientific Software with no distribution limitations – If the software has a PID, it must be provided to OSTI with the metadata record. If a PID is not provided, OSTI will assign one.
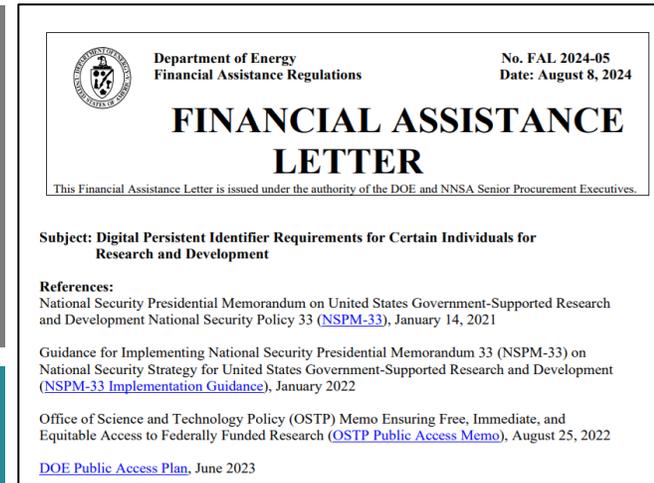
Requirements in effect – in a transition period for communities to adjust.

Persistent Identifier Implementation – PID FAL

NOFO: "Individuals required to submit Biographical Sketch and Current and/or Pending (Other) Support disclosures must provide a digital persistent identifier (PID) in such disclosures as part of the NOFO application. Included PIDs must meet the common/core standards specified in the NSPM-33 Implementation Guidance or successor guidance (e.g., an ORCID iD). The inclusion of an individual's PID will be optional until May 1, 2025, and mandatory thereafter."

Award Term: "Throughout the lifetime of the award, those individuals conducting research and development (R&D) under the award at the prime and subaward level must obtain and use a digital persistent identifier (PID) for themselves that meets the common/core standards specified in the NSPM-33 Implementation Guidance or successor guidance (e.g., an ORCID iD).

DOE requires recipients of federal awards with R&D activities, or technical assistance that supports R&D activities, to use the PID when publishing R&D outputs when that is an available option. Individuals conducting R&D activities at the prime and subaward level must report their R&D outputs as outlined in the DOE F 4600.2, U.S. Department of Energy "Federal Assistance Reporting Checklist" (FARC). The PID for individuals must be provided when reporting R&D outputs to the Department of Energy Office of Scientific and Technical Information (DOE OSTI)."

| Department of Energy | No. FAL 2024-05 |
| Financial Assistance Regulations | Date: August 8, 2024 |

FINANCIAL ASSISTANCE LETTER

This Financial Assistance Letter is issued under the authority of the DOE and NNSA Senior Procurement Executives.

Subject: Digital Persistent Identifier Requirements for Certain Individuals for Research and Development

References:
National Security Presidential Memorandum on United States Government-Supported Research and Development National Security Policy 33 (NSPM-33), January 14, 2021

Guidance for Implementing National Security Presidential Memorandum 33 (NSPM-33) on National Security Strategy for United States Government-Supported Research and Development (NSPM-33 Implementation Guidance), January 2022

Office of Science and Technology Policy (OSTP) Memo Ensuring Free, Immediate, and Equitable Access to Federally Funded Research (OSTP Public Access Memo), August 25, 2022

DOE Public Access Plan, June 2023

Requirement in effect for those who have included in NOFO and Award Term.

PID FAL also addresses PID requirements included in National Security Presidential Memo 33 (NSPM-33).

Financial Assistance Letter: *https://www.energy.gov/sites/default/files/2024-08/FAL24-05%20-%20PIDs%20for%20Individuals%20-Final%20080724.pdf*