Dataverse

Katie Mika
Data Services Librarian
Harvard Library & IQSS
katherine_mika@harvard.edu

Julian Gautier
Product Research Specialist
IQSS
juliangautier@g.harvard.edu

Gustavo Durand
Dataverse Technical Lead/Architect
IQSS
gdurand@g.harvard.edu

DATAVERSE.HARVARD.EDU
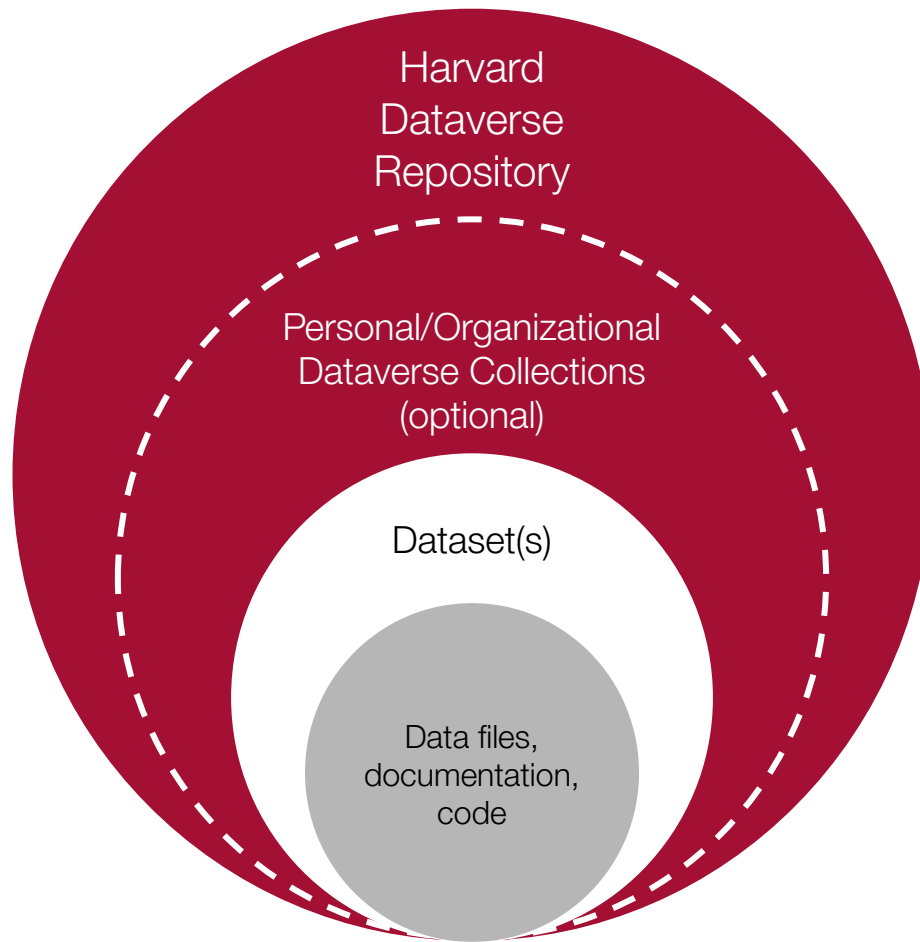
98 Installations

Map of Dataverse installations at https://dataverse.org/

Global Dataverse Community Consortium: https://dataversecommunity.global/

Harvard Dataverse Repository

Personal/Organizational Dataverse Collections (optional)

Dataset(s)

Data files, documentation, code

The Harvard Dataverse

- Free data repository open to all researchers from any discipline, both inside and outside of the Harvard community, where you can share, archive, cite, access, and explore research data.

- Collaboration between Harvard Library, Harvard University IT, and IQSS

- Self-curated repository in which users retain ownership over deposits

IQSS
The Institute for Quantitative Social Science

The Dataverse® Project

HARVARD LIBRARY   HARVARD UNIVERSITY Information Technology

FAIR Data

FINDABLE	ACCESSIBLE	INTEROPERABLE	REUSABLE
Well-described using standards Discoverable via aggregators & catalogs	Persistent identifier Citable Reliable storage	Standard & open formats, where possible Formats observe good preservation practices	Clear terms of use

Metadata	Technical Infrastructure	Policy & Rights

The FAIR Guiding Principles (Wilkinson et al. 2016)

IQSS
The Institute for Quantitative Social Science

The Dataverse Project

HARVARD LIBRARY

HARVARD UNIVERSITY
Information Technology

Who makes data FAIR?

Repositories

Assign persistent identifiers

Structure metadata records according to a disciplinary standard or schema

Index data as searchable resources

Retrieve datasets according to an open protocol that supports authentication

Preserves data files and metadata

Provenance and versions are tracked

Researchers

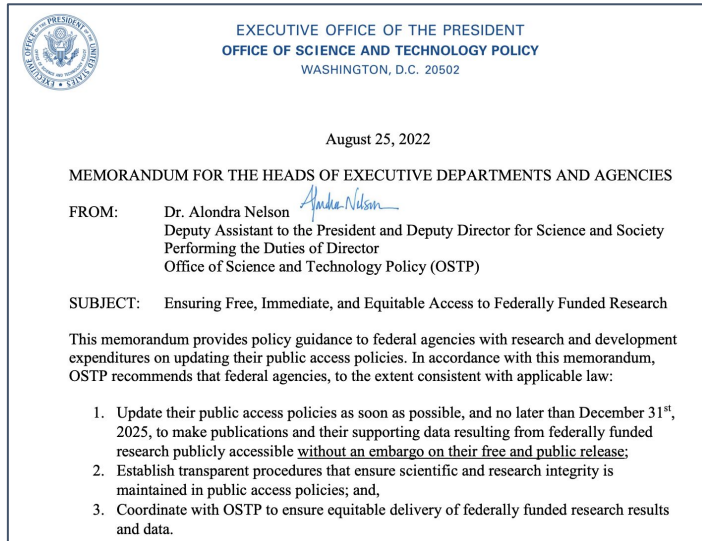Structure data clearly and apply good data management practices

Document data and software

Richly describe data using standardized metadata fields

Apply license and/or clear terms of use

Steward data across its usable life cycle

Response to Federal Mandates

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

August 25, 2022

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Dr. Alondra Nelson
Deputy Assistant to the President and Deputy Director for Science and Society
Performing the Duties of Director
Office of Science and Technology Policy (OSTP)

SUBJECT: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research

This memorandum provides policy guidance to federal agencies with research and development expenditures on updating their public access policies. In accordance with this memorandum, OSTP recommends that federal agencies, to the extent consistent with applicable law:

1. Update their public access policies as soon as possible, and no later than December 31st, 2025, to make publications and their supporting data resulting from federally funded research publicly accessible without an embargo on their free and public release;
2. Establish transparent procedures that ensure scientific and research integrity is maintained in public access policies; and,
3. Coordinate with OSTP to ensure equitable delivery of federally funded research results and data.

[OSTP Memo: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research](#)

Applications for Receipt Dates
ON/AFTER Jan 25 2023

NIH has issued the Data Management and Sharing (DMS) policy ⎘ (effective January 25, 2023) to promote the sharing of scientific data. Sharing scientific data accelerates biomedical research discovery, in part, by enabling validation of research results, providing accessibility to high-value datasets, and promoting data reuse for future research studies.

Under the DMS policy, NIH expects that investigators and institutions:

- Plan and budget for the managing and sharing of data
- Submit a DMS plan for review when applying for funding
- Comply with the approved DMS plan

Individual NIH Institutes, Centers, or Offices may have additional policies and expectations (see NIH Institute and Center Data Sharing Policies).

[NIH Data Management and Sharing Policy](#)

Current Dataverse Objectives

- Large data support
- Cross-domain metadata
- Computational workflows
- Harvesting metadata records
- Standardizing usage metrics
- Supporting sensitive data
- Improving UX & UI
- Curation services and training

Thanks!