

Librarians, funders, and the 2013 OSTP Public Access and Open Data Memoranda

Thea Atwood, Science and Engineering Library, University of Massachusetts Amherst, Amherst, MA

tpatwood@umass.edu

Summary

The NIH and the NSF are two of the largest funders of the sciences in the United States. This poster reports on how these two major funders have responded to two memoranda that focus on increasing access to the results of scientific research, and specific areas where librarians can provide assistance.

Background

In 2013, the Office of Science and Technology Policy released two memoranda:

Increasing Access to the Results of Federally Funded Scientific Research & Open Data Policy — Managing Information as an Asset (M-13-13)

Which directs Federal agencies with more than \$100M in R&D expenditures to:

- Make results of research freely available within one year of publication
- Require researchers to better manage and account for digital data
- Promote open standards and interoperability for digital data
- Establish a public data listing

The NSF released their plan in March of 2015, expecting compliance by January 2016.

The NIH released their plan in February of 2015, expecting compliance by the end of the calendar year 2015.

OSTP Guidelines	NIH	NSF
Publications		
Public access to peer-reviewed publications	PMC; multiple methods of compliance	NSF-PAR
Use a twelve-month embargo period	Twelve month maximum; Rights author can embargo for shorter period	Twelve months, and publications will require a permanent ID
Full access to publication metadata w/o charge	PubMed; ability to expand and deepen metadata	Minimum 8 metadata fields exposed
Use archival and accessibility best practices for	Archival XML	PDF/A; leverage other metadata solutions (CrossRef, ORCID, etc.)
Long-term preservation and access w/o charge	XML; Multiple backups internationally	PDF/A, dark archive w/ redundant backup
Use standard & nonproprietary file formats for publications and supplementary content	XML (rendered as PDF, HTML); PMC archives & makes public supplementary data	PDF/A; consult with community on data standards
Data		
Public access to digital data (with caveats for confidentiality, IP rights, etc.)	Supplementary data archived through PMC; no changes to reduce current protections; follow FAIR principles	Exploring how to best achieve public access, incl. preservation, access, & reuse
Balance demands of long-term preservation	Increased public access will inform guidance ; researchers' responsibility to balance	“Managing data is complex...”; look to communities for guidance; active research area
DMP included in all proposals	2003 Data Sharing Policy modified; will develop guidance	Since 2011; investigators report on progress annually & in final reports
Allow for funding to cover costs of data management	Data management considered a cost of conducting research	Can request funds to cover costs of making data available (GPG II.C.2.q.vi.(b))
Promote data deposited in public repositories	Supports numerous repositories; repositories should be first choice for dissemination	Follow publisher guidelines & DMP; Future: explore logistics of sharing ‘all’ data
Explore attribution to scientific data sets	Exploring ways of making data a “legitimate form of scholarship through data citation”	Can be included in biographical sketches
Improve data management practices & stewardship through training & education	Offers training initiatives, e.g., BD2K; will train staff and peer reviewers for DMP review	Funding through directorates; data science training via the Data Science Priority Goal
Open Data Policy		
Data created following best practices for long-term access and re-use	Referenced throughout policy; will apply to data generated under auspices of NIH	Only makes available agency data (e.g., NSF’s Facebook page)
Support data release and discovery through creation and maintenance of data inventory and public data listing	“Most digital scientific data...are also subject to expectations of the M-13-13 Open Data Policy.”; working with HHS to establish public data listing	“NSF does not collect data sets from recipients...so research data sets are not included in NSF’s public data listing.”
Adhere to security, privacy, and confidentiality	Considerations noted throughout the plan	Only agency-level data is released
Core agency processes adopt interoperability and openness requirements	Working with HHS to make administrative datasets available	Use open formats that can easily be shared via variety of mechanisms (e.g., Facebook, email)

Guidance for other funders

22 agencies have released public access plans.

More information on how each funder is responding to the OSTP’s memoranda is available at:

SPARC: <http://bit.ly/SPARC-PublicAccess>

ICPSR: <http://bit.ly/ICPSR-OSTP>

CENDI: <http://bit.ly/CENDI-PublicAccess>

Spreadsheet summarizing Federal public access policies: <http://bit.ly/FedOASummary>

How librarians can help:

- Compliance is up to institutions — look for opportunities to increase compliance (e.g., NIH Public Access Compliance Monitor)
- Help find ways to assign permanent identifiers
- Education on open access, open data, interoperability
- Education on data standards
- Data management and accounting for digital data
- Community guidance — still lacking in many areas; natural partnership with librarians
- Steer scholars away from storing supplemental and supporting data with publishers
- Facilitating the data stewardship, discussion of roles and responsibilities
- Providing researchers with realistic expectations for data curation work — i.e., data curation requires ongoing commitment
- Try to get on training grants, e.g., NSF-NRT
- Difference b/t sharing and usability (i.e., sharing something is not the same as making it usable)