

Preparing for new Federal guidelines on data publishing: Community readiness and Data records for NSF PAR

*Data Curation Network
September 21, 2021*

Overview

- EAGER project on Public Access and data records (and code)
- US Public Access policy and where it's headed
- NSF Public Access Repository
- Discussion – how do we better link science production with low-barrier public access reporting?
 - What are the biggest gaps in practices, systems and services that would deter or ease records submission to Federal Public Access Repositories?
 - Getting beyond the F and the A (in FAIR)...
 - AI-readiness (data / community practices)

EAGER: Community Readiness for Increasing Public Access to Data

- Initial work -
 - Gathering expert perspectives from leaders in scholarly communication and data infrastructure ;
 - Semi-structured interviews with domain researchers
 - Focus on Earth & Ag sciences (e.g. Soils; agronomy; other Earth Sci) and Data Science
 - particularly interest in projects using AI / ML methods, or producing data for reuse
 - Am actively recruiting for interviews – interest in early career, and “Rapid” awards
 - Engagement with a range of research communities on the use of the NSF Public Access Repository.

This work anticipates upcoming policy changes from NSF on the deposit of metadata for datasets resulting from funding (into the PAR).

US Public Access Policy

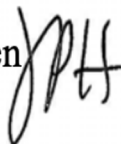
EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM:

John P. Holdren
Director

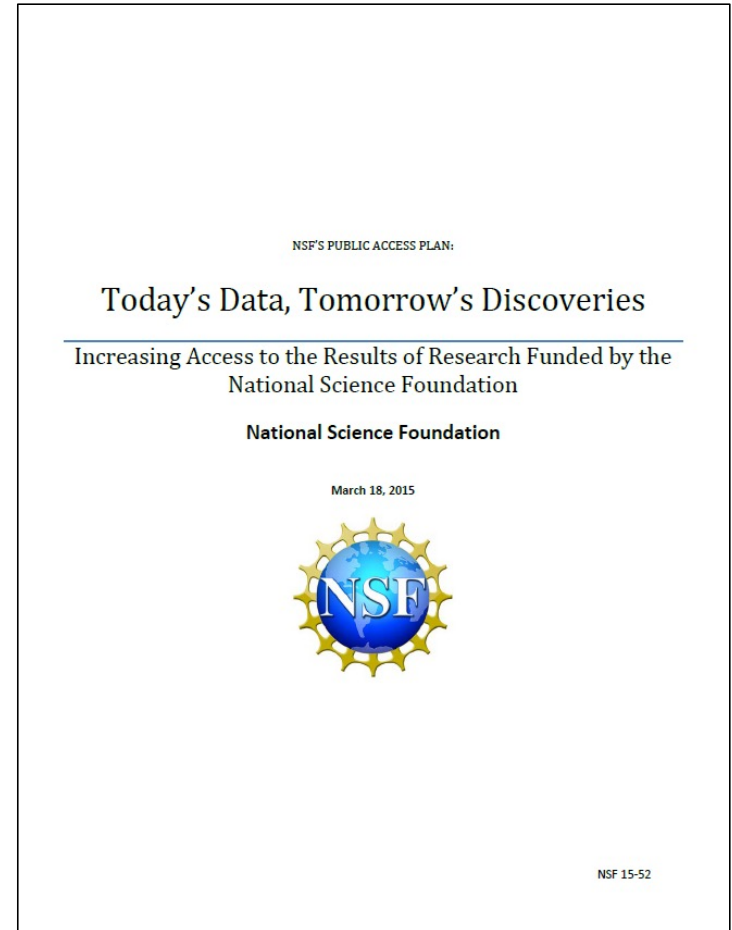


SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

1. Policy Principles

NSF 15-52 Plan

- NSF Public Access Initiative (PAI) was created by the NSF 15-52 Plan as the agency response to the 2013 U.S. Holdren Memorandum
- The Public Access Initiative is an NSF agency wide commitment, implemented through
 - policy,
 - procedures
 - infrastructure
- NSF holds a uniquely broad and influential role among organizational leaders of the commitment to public access in the United States



NSF Public Access Repository Functions and Characteristics – 1.0

PAR 1.0 is focused on Peer-Reviewed Articles

- Enables researchers to enter metadata for peer-reviewed articles or auto-populate by means of Digital Object Identifiers (DOI)
- Metadata recorded in PAR is also transmitted and synchronized with Award Search database
- The metadata for articles can be searched and displayed
- Researchers may deposit (or retrieve) public access articles as PDF/A files or point to download locations through DOI
- Note that PAR is not a single system, but rather an infrastructure of various interacting software systems in several different parts of the NSF infrastructure, also making use of modified modules from the DOE Office of Scientific and Technical Information (OSTI) system



<https://par.nsf.gov>

* Journal Name -OR- * Conference Proceeding Title

[Select journal name / ISSN](#)

ISSN: [?](#)

Volume:

Issue:

Page Range or eLocation-ID:

(e.g., 28 - 32)

Enter full date or year:

*Publication / Issue Full Date:



-OR-

*Publication / Issue Year:

(YYYY)

Time Period:

Separate multiple authors with a semicolon ;

*Authors: [?](#) Last name, First initial or name (middle name optional) e.g., Smith, John A; Anderson, S.

Separate multiple editors with a semicolon ;

*Editors: [?](#) Last name, First initial or name (middle name optional) e.g., Smith, John A; Anderson, S.

Description / Abstract:

NSF Public Access Repository (NSF-PAR)

A partnership with the Department of Energy,
Office of Scientific and Technical Information



Manage Publications

Contact

FAQ

Welcome, Cragin

Deposit Publication

1. Retrieve Publication Info

2. Deposit Final Accepted Version

3. Select Award & Acknowledge

4. Review

Enter the DOI number and click 'Submit' to retrieve the publication information from the publisher. Once you have reviewed the information, click 'Next' to proceed.

* Required Fields

* Enter Digital Object Identifier (DOI) Number:

Clear

[Where do I find my DOI number?](#)

Have you previously deposited your publication in another federal repository? [More Information](#)

Yes

No

Submit

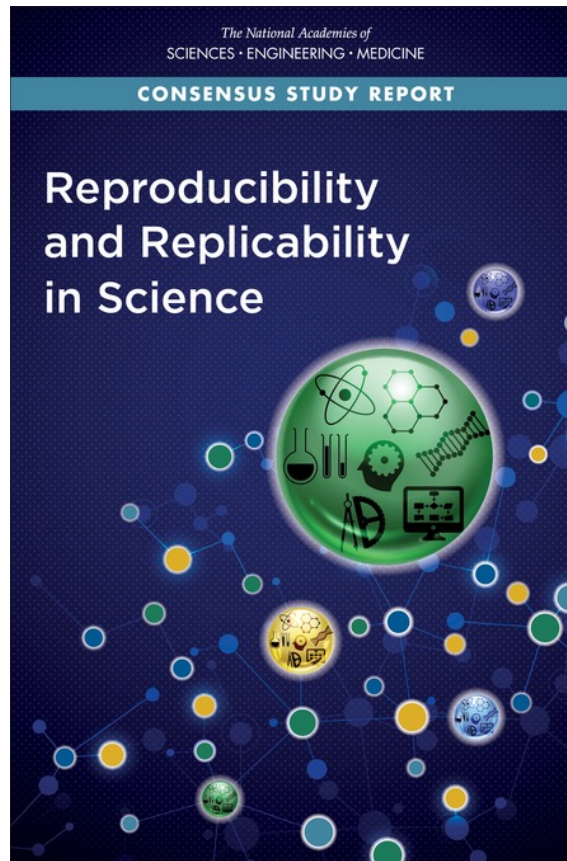
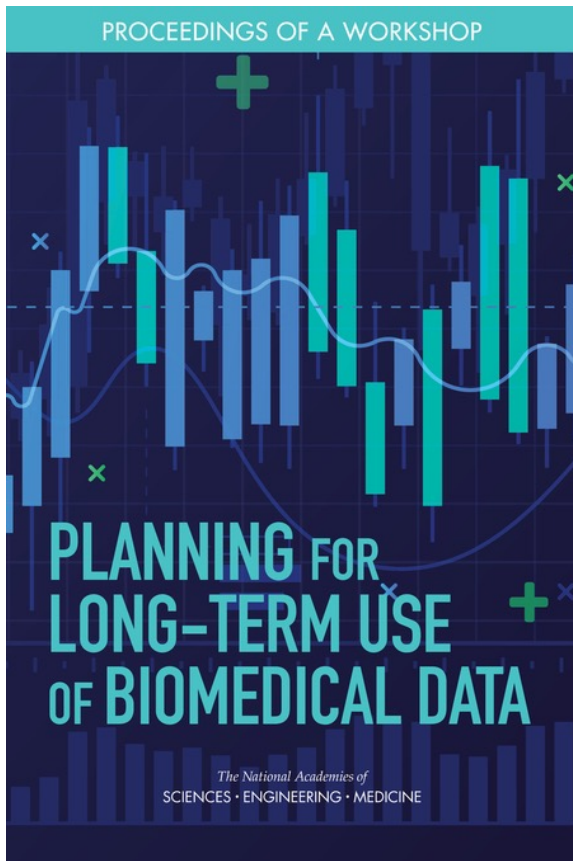
< Previous

Next >

What's coming for Data?



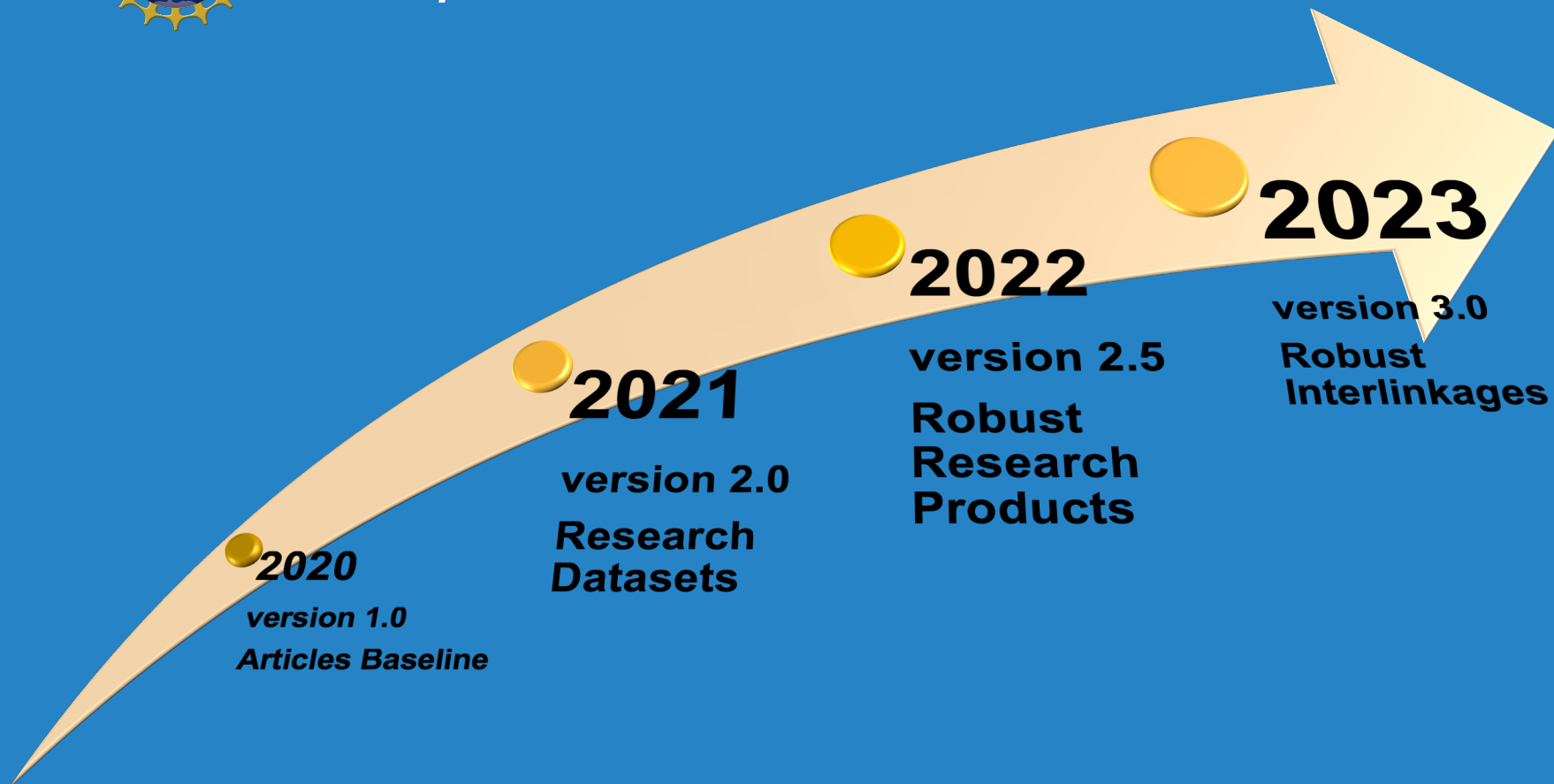
US Data Policy, Open Science





NSF Public Access Repository


Development Arc



M. Halbert, NSF - Used with permission

NSF Public Access Initiative



 [View Image Credit & Caption](#)

<https://beta.demo.nsf.gov/public-access>

[Home](#) / NSF Public Access Initiative

The NSF Public Access Initiative (PAI) strives to make the outputs of scientific research funded by the National Science Foundation publicly available to the greatest extent and with the fewest constraints possible and consistent with law.

- [NSF Public Access Repository \(NSF PAR\)](#)
- [See a list of what the PAI has funded](#)
- [FAQ for Public Access](#)

Questions / discussion

NSF PAR

- What kinds of information (or services) are needed to help researchers
 - deposit data in “preferred” repositories?
 - Have a better sense of the Public Access reporting process (submit records of these datasets)?

Adapting practices for publishing and Open Science

- Where (process) are the biggest barriers to researchers generating metadata records for datasets / code?
- Are publishers starting to monetize curation and stewardship?
- Where are we with AI-readiness?
 - machine-readable, AI-ready data – this is only part of the challenge
- What is the most pressing challenge that would benefit from closer collaboration with international efforts?

Thank you

mcragin@sdsc.edu

This work supported in part by NSF award 2032705.