



NATIONAL SCIENCE FOUNDATION
2415 EISENHOWER AVENUE
ALEXANDRIA, VIRGINIA 22314

NSF 20-068

Dear Colleague Letter: Open Science for Research Data

March 27, 2020

Dear Colleague:

Open science fuels scientific discovery and economic gain by making the products of Federally funded research more easily accessible and usable. Open science can also improve scientific rigor by directly linking the products of research (data and software) to their associated publications, making it easier for others to confirm the validity of a scientific result reported in a journal or juried conference proceeding.

In alignment with the benefits of open science, NSF is under taking an expansion of its Public Access Repository (NSF PAR) to include metadata records about the research data¹ that supports the journal and juried conference proceeding manuscripts resulting from NSF-funded research. The metadata records about the research data will contain sufficient information to allow for data discovery and an access determination to be made (but not all the metadata necessary for reuse of the research data). Research data will have a Digital Object Identifier (DOI)² that was assigned to it prior to it being reported to NSF. The research data will not reside in the NSF PAR but will instead reside in a repository, data center, or data portal managed by an organization that is committed to ensuring the availability of the data over time. The anticipated location of research data associated with a publication, if known, can be identified in the Data Management Plan and budgeted in the proposal.

Research data in support of a publication are i) the data necessary to confirm the validity of the scientific result reported in the publication, ii) the data described by the publication, or iii) as specified by the journal or conference proceeding.

Complementing the publication, the metadata record about research data in support of a publication will, as does the publication, become part of the public record on the NSF web site of the scientific contributions of an award. This extension to NSF PAR does not change the timing of reporting. Reporting is still done on an annual and final report basis, and at this time researchers will report on the products of their research that include both publications and supporting data. Data reporting will initially be voluntary.

Through this Dear Colleague Letter (DCL), the National Science Foundation (NSF) announces its intention to support conference proposals and EAGER proposals that explore and grow community readiness across all disciplinary areas served by the Foundation for this important advancement in open science as follows:

Proposals for Conferences: These are community workshops and other events that bring together stakeholders to explore and advance scientific community readiness in response to this advancement in open science.

Proposals for Early-Concept Grants for Exploratory Research (EAGER): These are for high-risk/high-reward innovative concepts and pilot project proposals that contribute to community readiness in response to this advancement in open science.

SUBMISSION PROCESS

Conference requests must not exceed \$50,000 in total for a one- or two-year duration. EAGERS can be supported up to \$300,000 for up to a two-year duration.

Inquiries about the applicability and relevance of an idea to the public access program must be made in advance, and NSF's response to this inquiry must be included in the proposal as a supplementary document. Inquiries will be accepted from a Principal Investigator (PI) or any consortium of investigators led by a PI at an eligible U.S. organization. An individual may serve as PI or co-PI on at most one conference proposal and one EAGER proposal pursuant to this DCL. Inquiries must be directed to the public access program director, Beth Plale, at bplale@nsf.gov. Proposals for awards made in fiscal year 2020 are due May 1, 2020.

Proposals must be prepared and submitted via Fastlane or Grants.gov, following the applicable instructions in the [NSF Proposal & Award Policies & Procedures Guide \(PAPPG\)](#). Guidance for EAGER Proposals is contained in PAPPG Chapter II.E.2 and guidance for Conference Proposals is contained in PAPPG Chapter II.E.7.

Proposals that fail to address the objectives and guidance described in this DCL will be returned without review.

Inquiries and questions should be directed to Beth Plale at bplale@nsf.gov.

Sincerely,

Joanne S. Tornow, Assistant Director, Biological Sciences
Margaret Martonosi, Assistant Director, Computer and Information Science and Engineering
Karen Marrongelle, Assistant Director, Education and Human Resources
Dawn M. Tilbury, Assistant Director, Engineering

William E. Easterling, Assistant Director, Geosciences
Anne L. Kinney, Assistant Director, Mathematical and Physical Sciences
Arthur W. Lupia, Assistant Director, Social, Behavioral, and Economic Sciences
Suzanne Iacono, Office Head, Office of Integrative Activities
Rebecca Keiser, Office Head, Office of International Science and Engineering

¹ According to OMB Circular A-81, section 200.315, "Research data means the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues." See <https://www.federalregister.gov/documents/2013/12/26/2013-30465/uniform-administrative-requirements-cost-principles-and-audit-requirements-for-federal-awards#sec-200-315>.

² Digital Object Identifier (DOI) is a globally unique Persistent ID (PID) that references research data. Research data can be referenced by more than one unique Persistent ID, thus can have a DOI along with having, for instance, an RRID, Handle, Clean URL, or proprietary registry number.