U.S. NATIONAL SCIENCE FOUNDATION
2415 EISENHOWER AVENUE
ALEXANDRIA, VIRGINIA 22314

NSF 24-085

Dear Colleague Letter: Leveraging Cyberinfrastructure for Research Data Management (RDM)

May 1, 2024

Dear Colleagues:

With this Dear Colleague Letter (DCL), the U.S. National Science Foundation (NSF), led by the Directorate for Computer and Information Science and Engineering (CISE) Office of Advanced Cyberinfrastructure (OAC), invites Early-Concept Grants for Exploratory Research (EAGER) and conference/workshop proposals that aim to leverage cyberinfrastructure to advance research data management (RDM) and public access to research data in alignment with the goals of OAC's Cyberinfrastructure for Public Access and Open Science (CI PAOS) program [https://new.nsf.gov/funding/opportunities/cyberinfrastructure-public-access-open-science-ci] and the NSF Public Access Initiative (PAI) [https://new.nsf.gov/public-access]. The NSF PAI aims to make the results of NSF-supported research publicly available to the greatest extent possible. The CI PAOS program aims to catalyze new and transformative socio-technical partnerships supporting research data infrastructure ecosystems across domains through early-stage collaborative activities between cyberinfrastructure researchers, scientists, research computing experts, data management experts, research labs, university libraries, and other communities of practice.

Advancements in data management across all research disciplines are needed to achieve national public access goals, including to "make publications and their supporting data resulting from federally funded research publicly accessible without embargo on their free and public release"[1] and implementing guidance to digital repositories that hold data from federally funded projects. This DCL responds to these needs while building upon previous NSF Dear Colleague Letters: Reproducibility and Replicability in Science (NSF 23-018), Open Science for Research Data (NSF 20-068), and Effective Practices for Data (NSF 19-069).

Accordingly, proposals are invited from across communities for Early-concept Grants for Exploratory Research (EAGER) projects and for workshops to address competency building,

capability building, and community building supporting RDM and public access. There is also a growing demand for international FAIR research collaboration, coordination, and RDM capacity. NSF particularly welcomes proposals involving new and innovative partnerships (such as between cyberinfrastructure (CI) researchers and providers, computer scientist (CS) and information technology (IT) researchers, domain scientists, and/or university libraries) that address RDM topics such as, but not limited to, the following examples:

1. Addressing data formats, metadata standards, and harmonization across domains while leveraging existing cyberinfrastructure to advance research data lifecycle management;
2. Addressing data curation, provenance, record keeping, organization of data, storage of data, linking, reporting of research data and results, sharing and access, institutional policies, and data management standard operating procedures for projects;
3. Promoting machine-readability, reproducibility, and utility for AI computing, resource, and tools to deploy modular RDM training for PIs, trainees, students at any scale;
4. Improving domain-specific data quality guidance, formats, RDM best practices, scientific connections, restrictive/sensitive RDM, transfer of large datasets, and case studies;
5. Leveraging existing data repositories or hardware/software computational resources for improving research data management, linking datasets within repositories, and combining data to gain new insights for sharing information with research communities;
6. Enabling community-scale integration of, and access to, data sets from funded research and/or those generated by industry to enable new research and discovery broadly for the CISE and other research communities;
7. Investigating interdisciplinary and international collaborations, costs, and continuity of RDM infrastructure for publications and data across projects, centers, and facilities for NSF-funded research;
8. Coordinating, facilitating, and/or evaluating U.S. and international FAIROS integrated research infrastructure efforts.

Proposals must follow the guidance contained in the NSF Proposal and Award Policies and Procedures Guide (PAPPG) for EAGER (see PAPPG Chapter II.F.3) and Conference/Workshop (see PAPPG Chapter II.F.8) proposals. Proposals should clearly describe the intellectual merit, broader impact, and budget for the proposed effort.

- **Early-Concept Grants for Exploratory Research (EAGER)** proposals are encouraged for high-risk/high-reward innovative concepts and pilot and other research proposals that contribute to community readiness and advancements in RDM for public access and open science. EAGERs can be supported up to $300,000 for up to a two-year duration.
- **Conference/Workshop** proposals are encouraged for community workshops and other events that bring together communities of practice to explore and advance scientific community readiness in response to this DCL (i.e., such as HPC experts, information

technology and computer scientists, library and information science, research labs, research libraries). Each Conference/Workshop proposal must not exceed $100,000 in total budget for a one-or two-year duration.

HOW TO RESPOND TO THIS DEAR COLLEAGUE LETTER

1. Proposers are required to email a Concept Outline, following the guidance in PAPPG Chapter I.D.1, to the cognizant Program Director listed in this DCL. The email subject header should start with "NSF DCL: Leveraging Cyberinfrastructure for Research Data Management (RDM)".
2. The PD may invite submission of an EAGER or Conference proposal based on an assessment of the alignment of the project description with the goals of the DCL and CI PAOS program [https://new.nsf.gov/funding/opportunities/cyberinfrastructure-public-access-open-science-ci].

When submitting your proposal to the CI PAOS Program Description, the project title should be of the form "CI PAOS: Title", after the PAPPG required prefix "EAGER" or "Conference". Please also include "Public Access" as a keyword in the at the beginning of the overview section of the project summary, and ensure inclusion of the following information in the project description:

- A dedicated section describing the Appropriateness of Proposed Work for EAGER or Conference funding,
- A dedicated section describing the relevance of the proposed work to CI PAOS goals (not to exceed one page).

After submitting the proposal, the NSF-assigned proposal number should be emailed to the program director identified below with the subject header starting with "NSF DCL: Leveraging Cyberinfrastructure for Research Data Management (RDM) - Proposal Number".

To be considered for priority consideration for FY 2024 funding, invited proposals must be submitted by 5:00 P.M. submitting organization's local time on July 3, 2024. There is not a deadline for concept outlines.

Concept outlines, as well as questions and inquiries about project ideas, should be sent to Plato Smith, Program Director, CISE/OAC, telephone: (703) 292-4278, email: plsmith@nsf.gov with a copy to FAIROSQueries@nsf.gov.

Sincerely,

Dilma Da Silva
Acting Assistant Director, Directorate for Computer and Information Science and Engineering (CISE)

Alexandra R. Isern
Assistant Director, Directorate for Geosciences (GEO)

Sylvia M. Butterfield
Acting Assistant Director, Directorate for Social, Behavioral and Economic Sciences (SBE)

C. Denise Caldwell
Acting Assistant Director, Directorate for Mathematical and Physical Sciences (MPS)

James L. Moore III
Assistant Director, Directorate for STEM Education (EDU)

Susan S. Margulies
Assistant Director, Directorate for Engineering (ENG)

Susan Marqusee
Assistant Director, Directorate for Biological Sciences (BIO)

Alicia J. Knoedler
Office Head, Office of Integrative Activities (OIA)

Kendra Sharp
Office Head, Office of International Science and Engineering (OISE)

---

[1] 2022 OSTP Memorandum (2022 Nelson Memo). https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-access-Memo.pdf.