**National Science Foundation**
**4201 Wilson Boulevard**
**Arlington, Virginia 22230**

**NSF 14-059**

**Dear Colleague Letter - Supporting Scientific Discovery through Norms and Practices for Software and Data Citation and Attribution**

Date: April 11, 2014

National Science Foundation
Directorate for Social, Behavioral & Economic Sciences (SBE)
Division of Social and Economic Sciences (SES)
Directorate for Computer & Information Science & Engineering (CISE)
Division of Advanced Cyberinfrastructure (ACI)

Dear Colleague:

How scientific research is conducted across all science disciplines is changing. One important direction of change is toward more open science, often driven by projects in which the output is purely digital, i.e., software or data. Scientists and engineers who develop software and generate data for their research spend significant time in the initial development of software or data frameworks, where they focus on the instantiation of a new idea, the widespread use of some infrastructure, or the evaluation of concepts for a new standard. Despite the growing importance of data and software products the effort required for their production is neither recognized nor rewarded. At present there is a lack of well-developed metrics with which to assess the impact and quality of scientific software and data. Unlike generally accepted citation-based metrics for papers, software and data citations are not systematically collected or reported. NSF seeks to explore new norms and practices in the research community for software and data citation and attribution, so that data producers, software and tool developers, and data curators are credited for their contributions.

The Science of Science and Innovation Policy (SciSIP) program and Software Infrastructure for Sustained Innovation (SI2) program invite collaborative workshop and exploratory research (EAGER) proposals in the following areas:

**Citation and attribution:**

- Novel mechanisms for citation of software and datasets as distinct products of scholarship, promoting standards of academic credit and rigor for these cyberinfrastructure components
- Novel citation methods for new forms of publication and scientific expression so that researchers are able to ensure their work is citable, and others are able to discover and access it
- Citation patterns that include a role for citations (e.g. to value activities such as "data provider/curator" and/or "software tool provider" alongside "data analyzer" or "computational modeler"), which can help create a credit market for data and software sharing

**Metrics of impact:**

- Introduction of appropriate metrics that match the effort necessary for successful development and maintenance of scientific software and data frameworks
- Development and use of metrics that measure software and data framework usage and impact on

science, engineering and education

- Establishment of metrics that recognize open access policies and sharing
- Comparison of impact for publication of software and data in a citable form before paper publication, as advocated in initiatives such as SageCite (http://blogs.ukoln.ac.uk/sagecite/), versus current practices
- Creation of specific project metrics that assess and monitor effective availability and accessibility of software and data
- Identification of sources of information about researchers' productivity and impact
- Development of ways in which researchers' scientific activity can be automatically captured and validated

It is strongly recommended that proposals be collaborative, with investigators from both the social science and cyberinfrastructure communities, including but not limited to: economics, sociology, science of organizations, management science, software infrastructure, data infrastructure, science of team science, science of science and innovation policy, and science, technology, and society.

This is not a new competition or a new program; interested investigators should first discuss their ideas with one of the program directors listed in this letter and upon approval may then submit an EAGER proposal or a proposal for Conferences, Symposia, and Workshops to the Science of Science & Innovation Policy (SciSIP) program via PD 09-7626 or to the Software Infrastructure for Sustained Innovation program via NSF 14-520. Proposals will be accepted through June 10, 2014.

Numerous Dear Colleague Letters and workshops supported by the Directorate for Computer & Information Sciences and Engineering (CISE) and the Directorate for Social, Behavioral, & Economic Sciences (SBE) have drawn attention to significant areas of research where potential citation and attribution problems have been identified, including, but not limited to:

- Describing the Conduct of Science in the Information Age - http://www.nsf.gov/pubs/2012/nsf12010/nsf12010.jsp
- Creating New Cyber-enabled Data on Innovation in Organizations - http://www.nsf.gov/pubs/2009/nsf09036/nsf09036.jsp
- Assessing and Enhancing the Impact of Science R&D in the United States: Chemical Sciences - http://www.nsf.gov/pubs/2010/nsf10054/nsf10054.jsp
- Assessing the Impacts of Recent and On-going Changes in Federal Science Policy - http://www.nsf.gov/pubs/2013/nsf13104/nsf13104.jsp
- Building a Community of Practice II - Report on the Second AAAS-NSF SciSIP Workshop http://scienceofsciencepolicy.net/publication/building-community-practice-ii-report-second-aaas-nsf-scisip-workshop
- Interdisciplinary Collaboration in Innovative Science & Engineering Fields - http://csid.unt.edu/nsf/idworkshop.php

The CISE Division of Advanced Cyberinfrastructure (ACI) and the SBE Division of Social and Economic Sciences (SES) encourage members of their research communities to consider this opportunity, either by participating directly in the submission of proposals, or by responding to invitations to participate in activities that may be funded through this Dear Colleague Letter.

Please refer to the following NSF documents for more information on ACI and SES priorities and interests:

- Cyberinfrastructure Framework for 21st Century Science and Engineering (CIF21) Vision - http://www.nsf.gov/cise/aci/cif21/CIF21Vision2012current.pdf
- ACCI Task Force on Software for Science and Engineering Report - http://www.nsf.gov/cise/aci/taskforces/TaskForceReport_Software.pdf
- A Vision and Strategy for Software for Science, Engineering, and Education -

http://www.nsf.gov/pubs/2012/nsf12113/nsf12113.pdf
- ACCI Task Force on Data and Visualization Report - http://www.nsf.gov/cise/aci/taskforces/TaskForceReport_Data.pdf
- A Vision and Strategy for Data in Science, Engineering, and Education - https://www.nsf.gov/cise/aci/cif21/DataVision2012.pdf
- Rebuilding the Mosaic Report - http://www.nsf.gov/pubs/2011/nsf11086/nsf11086.pdf

For further information on grant mechanisms, refer to the following documents:

- Early Concept Grants for Exploratory Research (EAGER) - http://www.nsf.gov/pubs/policydocs/pappguide/nsf14001/gpg_2.jsp#IID2
- Proposals for Conferences, Symposia and Workshops - http://www.nsf.gov/pubs/policydocs/pappguide/nsf14001/gpg_2.jsp#IID8


Please contact the following program directors prior to proposal submission and for any questions:

- Daniel Katz - CISE/ACI (dkatz@nsf.gov)
- Joshua Rosenbloom - SBE/SES (jlrosenb@nsf.gov)

Sincerely,

Joanne Tornow
Acting Assistant Director
Directorate for Social, Behavioral & Economic Sciences

Farnam Jahanian
Assistant Director
Directorate for Computer & Information Science & Engineering