Critical Techniques and Technologies for Advancing Foundations and Applications of Big Data Science & Engineering (BIGDATA)

PROGRAM SOLICITATION
NSF 15-544

REPLACES DOCUMENT(S):
NSF 14-543

National Science Foundation
Directorate for Computer & Information Science & Engineering
Directorate for Biological Sciences
Directorate for Education & Human Resources
Directorate for Engineering
Directorate for Geosciences
Directorate for Mathematical & Physical Sciences
Directorate for Social, Behavioral & Economic Sciences

Full Proposal Deadline(s) (due by 5 p.m. proposer's local time):
May 20, 2015

IMPORTANT INFORMATION AND REVISION NOTES

This revision of the BIGDATA solicitation differs from 14-543 in the following ways:

1. The solicitation includes participation by the Mathematical and Physical Sciences (MPS) and Geosciences (GEO) Directorates at NSF.
2. The Office of Financial Research, Department of Treasury, has joined the program.
3. Proposals must be submitted under one of two categories, Foundations (F) or Innovative Applications (IA), or under both. However, there are no further sub-categories based on research themes.
4. An individual may participate as PI, co-PI, Senior Personnel, consultant, or any other role in no more than one proposal, or related sub-award.
5. The page limit for Project Description is fifteen (15) pages.
6. A separate Project Coordination Plan, not exceeding two (2) pages, is required for all proposals, and should be included under Supplementary Documents.
7. The Data Management Plan can be up to three (3) pages.

Important Information

Any proposal submitted in response to this solicitation should be submitted in accordance with the revised NSF Proposal & Award Policies & Procedures Guide (PAPPG) (NSF 15-1), which is effective for proposals submitted, or due, on or after December 26, 2014. The PAPPG is consistent with, and, implements the new Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards (Uniform Guidance) (2 CFR § 200).

SUMMARY OF PROGRAM REQUIREMENTS

General Information

Program Title:
Critical Techniques and Technologies for Advancing Foundations and Applications of Big Data Science & Engineering (BIGDATA)

Synopsis of Program:
The BIGDATA program seeks novel approaches in computer science, statistics, computational science, and mathematics, along with innovative applications in domain science, including social and behavioral sciences, geosciences, education, biology, the physical sciences, and engineering that lead towards the further development of the interdisciplinary field of data science. The solicitation invites two types of proposals: "Foundations" (F): those
developing or studying fundamental theories, techniques, methodologies, technologies of broad applicability to Big Data problems; and “Innovative Applications” (IA): those developing techniques, methodologies and technologies of key importance to a Big Data problem directly impacting at least one specific application. Therefore, projects in this category must be collaborative, involving researchers from domain disciplines and one or more methodological disciplines, e.g., computer science, statistics, mathematics, simulation and modeling, etc. While Innovative Applications (IA) proposals may address critical big data challenges within a specific domain, a high level of innovation is expected in all proposals and proposals should, in general, strive to provide solutions with potential for a broader impact on data science and its applications. IA proposals may focus on novel theoretical analysis and/or on experimental evaluation of techniques and methodologies within a specific domain. Proposals in all areas of sciences and engineering covered by participating directorates at NSF are welcome.

While notions of volume, velocity, and variety are commonly ascribed to big data problems, other key issues include data quality and provenance. Data-driven solutions must carefully ascribe quality and provenance to results in a manner that is helpful to the users of the results. For example, in some cases, such as in education research, data quality may aggregate to test or measurement instrument quality, where a composite of variables may be used to describe one or more constructs.

In addition to approaches such as search, query processing, and analysis, visualization techniques will also become critical across many stages of big data use—to obtain an initial assessment of data as well as through subsequent stages of scientific discovery. Research on visualization techniques and models will be necessary for serving not only the experts, who are collecting the data, but also those who are users of the data, including “cross-over” scientists who may be working with big data and analytics for the first time, and those using the data for teaching at the undergraduate and graduate levels. The BIGDATA program seeks novel approaches related to all of these areas of study.

Before preparing a proposal in response to this BIGDATA solicitation, applicants are strongly urged to consult the list of related solicitations available at: http://www.nsf.gov/cise/news/bigdata.jsp and consult the respective NSF program officers listed in them should those solicitations be more appropriate. In particular, applicants interested in deployable cyberinfrastructure pilots that would support a broader research community should see the Campus Cyberinfrastructure - Data, Networking, and Innovation Program (CC/DNI) solicitation (http://www.nsf.gov/pubs/2015/nsf15534/nsf15534.htm?WT.mc_id=USNSF_25&WT.mc_ev=click). Applicants should also consider the Computational and Data Enabled Science and Engineering (CDS&E, PD 12-8084) (http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504813) and Exploiting Parallelism and Scalability (XPS, NSF 15-511) (http://www.nsf.gov/pubs/2015/nsf15511/nsf15511.htm) solicitations for potential fit.

Cognizant Program Officer(s):

Please note that the following information is current at the time of publishing. See program website for any updates to the points of contact.

- Chaitanya Baru, Senior Advisor - CISE/OAD, telephone: (703) 292-4541, email: cbaru@nsf.gov
- Sylvia Spengler, Program Director - CISE/IIS, telephone: (703) 292-8930, email: sspengle@nsf.gov
- Jack Snoeyink, Program Director - CISE/CCF, 1115, telephone: (703) 292-8910, email: j snoeyin@nsf.gov
- Amy Apon, Program Director - CISE/CNS, telephone: 703-292-8950, email: aapon@nsf.gov
- Elizabeth R. Blood, Program Director - BIO/OAD, telephone: (703) 292-4349, email: eblood@nsf.gov
- Helen T. Martin, Program Director- EHR/DRL, telephone: (703) 292-7016, email: htmartin@nsf.gov
- George Haddad, Program Director - ENG/ECCS, telephone: (703) 292-8897, email: ghaddad@nsf.gov
- Mona Zaghliou, Program Director - NEG/ECCS, telephone: (703) 292-8339, email: mzaghliou@nsf.gov
- Eva Zanzerka, Program Director - GEO/OAD, telephone: (703) 292-4734, email: ezanzerk@nsf.gov
- Xiaoming Huo, Program Director - MPS/DMS, telephone: (703) 292-4876, email: xhuo@nsf.gov
- Nandini Kannan, Program Director - MPS/DMS, telephone: (703) 292-8104, email: nakannan@nsf.gov
- Bogdan Mihaila, Program Director - MPS/PHY, telephone: (703) 292-8235, email: bmihaila@nsf.gov
- Heng Xu, Program Director - SBE/SES, telephone: (703) 292-8643, email: hxu@nsf.gov
- Edward Garnett, Program Manager -- OFR/Dept. of Treasury, telephone: (202) 927-8025, email: edward.garnett@treasury.gov

Applicable Catalog of Federal Domestic Assistance (CFDA) Number(s):

- 47.041 --- Engineering
- 47.049 --- Mathematical and Physical Sciences
- 47.050 --- Geosciences
- 47.070 --- Computer and Information Science and Engineering
- 47.074 --- Biological Sciences
- 47.075 --- Social Behavioral and Economic Sciences
- 47.076 --- Education and Human Resources

Award Information

Anticipated Type of Award: Standard Grant or Continuing Grant or Cooperative Agreement

Estimated Number of Awards: 27 to 35

About 27-35 projects will be funded, subject to availability of funds.
Anticipated Funding Amount: $26,500,000

Up to $26,500,000 will be invested in proposals submitted to this solicitation, subject to availability of funds. Projects will be funded in the range of $200,000 to a maximum of $500,000 per year in total funding for 3 to 4 years of support.

Eligibility Information

Who May Submit Proposals:
The categories of proposers eligible to submit proposals to the National Science Foundation are identified in the Grant Proposal Guide, Chapter I, Section E.

Who May Serve as PI:
There are no restrictions or limits.

Limit on Number of Proposals per Organization:
There are no restrictions or limits.

Limit on Number of Proposals per PI or Co-PI: 1

An individual may participate as PI, co-PI, Senior Personnel, consultant, or any other role in no more than one proposal, or related sub-award, submitted in response to this solicitation.

In the event that an individual exceeds this limit, any proposal submitted to this solicitation with this individual listed as a PI, co-PI, senior personnel, consultant or any other role after the first proposal is received at NSF will be returned without review. No exceptions will be made.

Proposals submitted in response to this solicitation may not duplicate or be substantially similar to other proposals concurrently under consideration by NSF.

Proposal Preparation and Submission Instructions

A. Proposal Preparation Instructions
   - Letters of Intent: Not required
   - Preliminary Proposal Submission: Not required
   - Full Proposals:

B. Budgetary Information
   - Cost Sharing Requirements: Inclusion of voluntary committed cost sharing is prohibited.
   - Indirect Cost (F&A) Limitations: Not Applicable
   - Other Budgetary Limitations: Not Applicable

C. Due Dates
   - Full Proposal Deadline(s) (due by 5 p.m. proposer's local time):
     May 20, 2015
     Full Proposal Due

Proposal Review Information Criteria

Merit Review Criteria: National Science Board approved criteria. Additional merit review considerations apply. Please see the full text of this solicitation for further information.

Award Administration Information

Award Conditions: Standard NSF award conditions apply.

Reporting Requirements: Standard NSF reporting requirements apply.

TABLE OF CONTENTS
I. INTRODUCTION

The Big Data phenomenon has emerged as a result of vast amounts of data that are becoming available across a wide range of application domains across science, business, and government. Today, scientists, engineers, educators, citizens and decision-makers have unprecedented amounts and types of data available to them. Data may originate from many disparate sources, including scientific instruments, medical devices, telescopes, microscopes, satellites; digitally-authored media, including text, images, audio, and emails; streaming data from weblogs, videos, financial/commercial transactions; from ubiquitous sensing and control applications in engineered and natural systems, through multitudes of heterogeneous sensors and controllers instrumenting these systems; social interactional data from social networking sites, twitter feeds and click streams; administrative data; or scientific data from large-scale surveys, brain research, large-scale simulations, continuous simulation models, and computational analyses of observational data. The data can be temporal, spatial, or dynamic, structured or unstructured; and the information and knowledge derived from data can differ in representation, complexity, granularity, context, quality, provenance, reliability, trustworthiness, and scope. Data collected on student learning may reflect task attributes (e.g., calculation routines), cognitive demands (i.e., understanding the basis for the calculation or its application), memory and executive control processes (e.g., recall, attention, decision making), and a variety of "non-cognitive" attributes (e.g., lack of persistence in problem solving).

Realizing the transformative potential of big data requires addressing many challenges in the management of data and knowledge; automation of many aspects of the data-enabled discovery processes; the computational, mathematical, and statistical methods for data analysis; and visualization techniques for “making sense” of the data. Principled approaches are needed to address the management, modeling, and analysis of such data. The challenges in managing and analyzing the volumes and rates of newly available data can require fundamentally new techniques and technologies in order to handle the complexity, size, or rate of these data. Methods are needed for supporting data experiments: for browsing through the data; characterizing, summarizing and integrating data; and for hypothesis generation and testing. Attention should be placed on exploring the capacity to draw causal inferences when random assignment is not feasible or possible, and/or when structural or laboratory controls are unavailable. Issues related to multicollinearity and prediction, statistical vs. substantive significance, and other problems of inference should also be considered. Analytical techniques, including signal processing methods, can also be developed to provide novel ways to access and analyze newly generated data streams, or existing data. Design and development should ideally take place with a partnership model to assure that the new techniques and technologies will result in the ability to attain a fuller understanding of the data and their underlying context, and support the formulation and investigation of pertinent questions. While the context is centrally important, one of the challenges may be in characterizing and codifying the context.

While the advent of big data offers unprecedented opportunities for data-driven discovery and decision-making in virtually every area of human endeavor, ensuring appropriate privacy and security guarantees becomes increasingly challenging as well. Novel techniques may need to be investigated and developed for data confidentiality and privacy in data collection, storage, transfer, archiving, and secondary or third-party analyses. Also important are issues related to reproducibility, replicability and uncertainty quantification.

The Critical Techniques and Technologies for Advancing Foundations and Applications of Big Data Science & Engineering (BIGDATA) program invites proposals in two categories:

- "Foundations" (F): those that focus on the development of novel techniques, or novel theoretical analysis (including statistics and probability) or experimental evaluation of techniques, that are broadly applicable; and
- "Innovative Applications" (IA): those that focus on the development of innovative techniques, methodologies, and technologies for specific application areas or innovative adaptations of existing techniques, methodologies and technologies to new application areas.

This solicitation is one component of the "National Big Data Initiative", a coordinated set of activities involving multiple federal agencies. The Big Data Initiative aims to make the most of the fast-growing volume of digital data by improving the ability to extract knowledge and insights from large and complex collections of digital data to help solve some the Nation’s most pressing challenges. The Initiative strives to:

- Advance state-of-the-art core technologies needed to collect, store, preserve, manage, analyze, and share huge quantities of data.
- Harness these technologies to accelerate the pace of discovery in science and engineering, strengthen our national...
Proposals submitted in response to the BIGDATA solicitation must focus on the development of the scientific discovery processes, novel systems, novel computational, statistical, or mathematical techniques and technologies, or novel theoretical analyses or experimental evaluations of techniques for big data and knowledge management and analytics. BIGDATA proposals submitted under the “Foundations” (F) category must be highly innovative and broadly applicable. BIGDATA proposals submitted under the "Innovative Applications" (IA) category must address a big data challenge of key importance to at least one application domain from one of the participating NSF directorates, and involve a substantial collaboration between researchers in the application domain and computational disciplines. When possible, submitted proposals are expected to support efforts for the preparation of scientists to confront new challenges in data science, through training and education activities, and research on learning how to model, understand and communicate insights possible in big data.

Foundations: Proposals submitted to the "Foundations" (F) category are expected to address the development of highly innovative fundamental techniques, theories, methodologies, and technologies for big data management or analytics, including knowledge management and semantic technologies; or, novel analyses of existing techniques and methodologies, whose solutions have wide applicability beyond specific narrow domains. Proposals that focus only on the scaling up of existing methods should not be submitted under this category unless the technique used to scale the existing method is itself highly innovative. Assembly and analysis of specific datasets may be part of these efforts, provided that the purpose is the development and testing of algorithms, techniques, technologies, and methodologies, developed in these projects. Proposals aimed at the development and deployment of big data infrastructure can be responsive to this solicitation if the work proposed to accomplish these aims includes substantial innovation.

Innovative Applications: Proposals submitted to the "Innovative Applications" (IA) category must address a big data problem of key importance to at least one domain. Such problems require close interactions among computer scientists, mathematicians, statisticians, and domain scientists and engineers in order to address complex, data-driven questions, including development of domain knowledge structures and ontologies, in one or more domains. Early engagement with domain scientists/engineers provides a partnership design and development model to help focus on the right problems to solve, and to help assess the quality of the solution in a realistic setting, while providing a fuller understanding of the real constraints of the applications and data in the domains. Therefore, projects in this category must be collaborative, involving researchers from domain disciplines and one or more methodological disciplines, e.g., computer science, mathematics, statistics, simulation and modeling, etc., stimulating further research on all sides of the collaboration.

Proposals in this category must demonstrate a linkage to needs and challenges in one or more domains. The proposed research can range from technology transfer, including the scaling up of existing methods to analyze larger or more complex data, to the development of fundamentally new methods. However, innovation is an important criterion and proposals should identify the technical challenges involved in the development of the proposed methods. Research in this category is expected to be multi-faceted, and may include the development of a computational infrastructure or the assembly and analysis of large or complex datasets. However, the development of innovative new techniques, innovative computational methods and analyses, or improved understanding of existing techniques, must be a major goal of the project, and should be integrated with the other research goals of the project.

Proposals awarded in this category are expected to be substantially funded by the NSF directorate(s), or participating agencies (in this case, the Office of Financial Research, OFR), interested in the application area. Therefore, applicants considering submitting proposals in this category are strongly encouraged to discuss their planned research with a program officer from the directorate, or participating agency (OFR), in advance of submitting the proposal, and state the name of the directorate(s), or agency, on a separate line at the bottom of the broader impacts text box. (See instructions in section V.A.2, under Project Summary).

Proposals may involve both themes, i.e., Foundations as well as Innovative Applications. See section V.A.1 Proposal Titles, below, on how to identify such proposals.

II.A Office of Financial Research-specific Topics

The Office of Financial Research (OFR) is a participant in the current NSF BIGDATA solicitation. OFR has a mandate to research, identify and assess risks to the financial stability of the United States and to monitor, investigate, and report such risks to Congress and the Financial Stability Oversight Council ("Council"). Consequently, OFR's primary collaborative interest is in the area of Financial Research and Data Analysis. OFR and NSF have shared interests at the intersection of these two areas, centered around computational and information processing approaches as referenced below.

II.A.1 OFR Topic Areas of Interest

Topics of research interest to OFR include, but are not limited to: Analysis of financial networks; Algorithms and methods for measuring threats to financial stability; Representation and standardization of financial data and information; Formal methods for representation and analysis of financial contracts and regulations, e.g., logics, ontologies, and rule-based approaches; Complexity of financial systems and relationships; Technologies for modeling and monitoring financial systems and infrastructure; Aggregation methods for complex financial data; Visualizations of the financial system and its attendant risks; Financial risk management techniques for the quantification of uncertainty and risk, including stress testing, risk and volatility forecasting, and the modeling of statistical distributions, processes, and dependence structures; Representation and querying of uncertain financial data, such as marks to model for infrequently traded instruments; Storage and query tools and techniques applicable to financial data; Assembly, integration, and analysis of new datasets for financial research; Techniques for ensuring the security and confidentiality of sensitive financial data, including approaches for selective sharing; Technologies and methodologies to support investigations of failures and disruptions in financial markets, such as those that might arise from (or be exacerbated by) automated high frequency trading systems; Simulation of financial systems, for example using Monte Carlo and agent-based methods; and Tools to support financial policymaking and decision-making.

III. AWARD INFORMATION
Estimated program budget, number of awards and average award size/duration are subject to the availability of funds. An estimated 27-35 projects will be funded, subject to availability of funds. Up to $26,500,000 will be invested in proposals submitted to this solicitation, subject to availability of funds.

Award sizes can range from $200,000 to $500,000 per year for three to four years in total funding.

IV. ELIGIBILITY INFORMATION

Who May Submit Proposals:

The categories of proposers eligible to submit proposals to the National Science Foundation are identified in the Grant Proposal Guide, Chapter I, Section E.

Who May Serve as PI:

There are no restrictions or limits.

Limit on Number of Proposals per Organization:

There are no restrictions or limits.

Limit on Number of Proposals per PI or Co-PI: 1

An individual may participate as PI, co-PI, Senior Personnel, consultant, or any other role in no more than one proposal, or related sub-award, submitted in response to this solicitation.

In the event that an individual exceeds this limit, any proposal submitted to this solicitation with this individual listed as a PI, co-PI, senior personnel, consultant or any other role after the first proposal is received at NSF will be returned without review. No exceptions will be made.

Proposals submitted in response to this solicitation may not duplicate or be substantially similar to other proposals concurrently under consideration by NSF.

Additional Eligibility Info:

V. PROPOSAL PREPARATION AND SUBMISSION INSTRUCTIONS

A. Proposal Preparation Instructions

Full Proposal Preparation Instructions: Proposers may opt to submit proposals in response to this Program Solicitation via Grants.gov or via the NSF FastLane system.

- Full proposals submitted via FastLane: Proposals submitted in response to this program solicitation should be prepared and submitted in accordance with the general guidelines contained in the NSF Grant Proposal Guide (GPG). The complete text of the GPG is available electronically on the NSF website at: http://www.nsf.gov/publications/pub_summ.jsp?ods_key=gpg. Paper copies of the GPG may be obtained from the NSF Publications Clearinghouse, telephone (703) 292-7827 or by e-mail from nsfpubs@nsf.gov. Proposers are reminded to identify this program solicitation number in the program solicitation block on the NSF Cover Sheet For Proposal to the National Science Foundation. Compliance with this requirement is critical to determining the relevant proposal processing guidelines. Failure to submit this information may delay processing.

- Full proposals submitted via Grants.gov: Proposals submitted in response to this program solicitation via Grants.gov should be prepared and submitted in accordance with the NSF Grants.gov Application Guide: A Guide for the Preparation and Submission of NSF Applications via Grants.gov. The complete text of the NSF Grants.gov Application Guide is available on the Grants.gov website and on the NSF website at: (http://www.nsf.gov/publications/pub_summ.jsp?ods_key=grantsgovguide). To obtain copies of the Application Guide and Application Forms Package, click on the Apply tab on the Grants.gov site, then click on the Apply Step 1: Download a Grant Application Package and Application Instructions link and enter the funding opportunity number, (the program solicitation number without the NSF prefix) and press the Download Package button. Paper copies of the Grants.gov Application Guide also may be obtained from the NSF Publications Clearinghouse, telephone (703) 292-7827 or by e-mail from nsfpubs@nsf.gov.

In determining which method to utilize in the electronic preparation and submission of the proposal, please note the following:

Collaborative Proposals. All collaborative proposals submitted as separate submissions from multiple organizations must be submitted via the NSF FastLane system. Chapter II, Section D.5 of the Grant Proposal Guide provides additional information on collaborative proposals.

See Chapter II.C.2 of the GPG for guidance on the required sections of a full research proposal submitted to NSF. Please note that the proposal preparation instructions provided in this program solicitation may deviate from the GPG instructions.

1. Proposal Titles: Proposal titles must indicate the BIGDATA program, followed by a colon, then the category (F for "Foundations", IA for "Innovative Applications"), followed by the title of the project. Thus, titles would be BIGDATA:F:Title or BIGDATA:IA:Title. If both the themes are involved, they should be ordered by their importance in the research being proposed. For example, BIGDATA:F:IA:Title, if foundational technology is the more important thrust of the proposal, but a specific domain is also included. Or, BIGDATA:IA:F:Title, if vice versa. Titles of collaborative proposals should be prepared in a similar manner, but should also include "Collaborative Research" immediately after Bigdata. Thus, the title of each proposal for a collaborative set of proposals will be BIGDATA:Collaborative Research..., followed by the proposal category(ies) and actual title.

2. Project Summary (1-page limit, 4,600 characters in total for all three text boxes): At the top of the Overview text box enter
the title of the BIGDATA project, the name of the PI and the lead institution. Provide a summary description of the BIGDATA project, including its transformative research and education goals, and the community (communities) that will be impacted by its results. In the separate text boxes, provide a succinct summary of the intellectual merit and broader impacts of the proposed project. Full proposals that do not address the intellectual merit and broader impacts of the proposed project in separate statements within the project summary will be returned without review.

Proposals submitted to the “Innovative Applications” (IA) category must specify one or more relevant participating NSF directorates, or agencies, for the project on a separate line at the bottom of the broader impacts text box. The list of participating NSF directorates includes Biological Sciences (BIO), Computer and Information Science and Engineering (CISE), Education and Human Resources (EHR), Engineering (ENG), Geosciences (GEO), Mathematical and Physical Sciences (MPS), and Social, Behavioral & Economic Sciences (SBE). The participating agency is the Office of Financial Research (OFR).

All project summaries must include a list of three to six keywords at the bottom of the broader impacts text box.

3. Project Description. There is a 15-page limit for the Project Description.

The project description should clearly identify the transformative potential of the proposed research, how it addresses a big data challenge, the applicability of the proposed technique or technology to a specific domain, or more broadly, and the novelty of the approach.

An evaluation plan is expected in any BIGDATA proposal, and should be provided within the Project Description. The evaluation plan should describe how the techniques, methodologies, theories, and technologies would be evaluated, and can include results from applications of that technology to specific domains, efficacy studies, and other such activities. Where applicable, evaluation plans should specify benchmark datasets and competing methods that will be used to evaluate scalability and performance, or address the development of new benchmarks, if needed. The evaluation plan should be appropriate for the size and scope of the project.

The project description must contain, as a separate section within the narrative, a section labeled “Broader Impacts of the Proposed Work”, as described in Chapter II.C.2.d of the NSF Grant Proposal Guide. Proposals that omit a separate section that deals explicitly with Broader Impacts will be returned without review.

4. Proposal Budget. It is expected that at least one PI from each funded project will attend an annual BIGDATA Principal Investigator (PI) meeting to present project research findings and broader impacts. Collaborative proposals involving investigators from multiple disciplines should have PIs from each of the major disciplines for the grant attending the annual meeting. Requested budgets should include funds for travel to this event.

5. Supplementary Documents. Supplementary documents are limited to the specific types of documentation listed in the GPG, with a few exceptions, as specified below. Proposals containing special information or supplementary documentation that has not been explicitly allowed in the GPG or this solicitation, such as article reprints or preprints, or appendices, will be returned without review.

Collaborative proposals. Simultaneously submitted collaborative proposals, and proposals that include subawards, are a single unified project; in many cases, supplementary documents for such proposals should only be provided in the proposal submitted by the lead institution. See below for specific instructions for each supplementary document type.

a. Project Coordination Plan - Required. Proposals must include a supplementary document of no more than two pages labeled “Project Coordination Plan”, which must include 1) the specific roles of the collaborating PIs, Co-PIs, other Senior Personnel and paid consultants at all organizations involved; 2) how the project will be managed across institutions and disciplines; 3) identification of the specific coordination mechanisms that will enable cross-institution and/or cross-discipline scientific efforts, e.g., webinars, workshops, software exchange, project meetings at conferences, use of videoconferencing and other communication tools, software repositories, etc.; 4) specific references to the budget line items that support these coordination mechanisms; and 5) a detailed timeline.

b. Data Management Plan, including Software Sharing - Required. Proposals must include a supplementary document of no more than three pages labeled “Data Management Plan”, which must include a Software Sharing Plan (see below). This supplement is provided in full by the lead institution. The data management plan should describe how the project will manage its data and software and share its research results (including software). The Data Management Plan will be evaluated using NSF review criteria. Data management requirements and plans specific to the Directorate, Office, Division, Program, or other NSF unit, relevant to a proposal are available at, http://www.nsf.gov/bfa/dias/policy/dmp.jsp.

The plan may include information on:

i. The types of data, software, curriculum materials, and other materials to be produced in the course of the project;

ii. The standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);

iii. Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;

iv. Policies and provisions for re-use, re-distribution, and the production of derivatives;

v. Plans for archiving data, samples, and other research products, and for preservation of access to them; and

vi. Software sharing plan, with appropriate timelines, as described below.

The Software Sharing Plan (with appropriate timelines) is required of all proposals as part of the Data Management Plan. There is no prescribed single license for software produced through grants responding to this announcement. However, the Program does have goals for software dissemination, and reviewers will be instructed to evaluate the dissemination plan relative to these goals:

1. The software should be freely available to science and engineering researchers and educators in the non-profit sector, such as institutions of education, research institutions, and government laboratories.

2. The terms of software availability should permit the dissemination and commercialization of enhanced or customized versions of the software, or incorporation of the software or pieces of it into other software packages.

3. To preserve utility to the community, the software should be transferable so that another individual or team can continue development in the event that the original investigators are unwilling or unable to do so.

4. The terms of software availability should include the ability of researchers to modify the source code and to share modifications with other colleagues. An applicant should take responsibility for creating the original and subsequent official versions of a piece of software.
5. To further enhance the potential impact of their software, applicants may consider proposing a plan to manage and disseminate the improvements or customizations of their tools and resources by others. This proposal may include a plan to incorporate the enhancements into the official core software, may involve the creation of an infrastructure for plug-ins, or may describe some other solution.

6. If a particular license is selected for the software distribution, it should be specified in the proposal. If an open-source license is not selected, the proposal should explain why this decision was made.

c. Human and vertebrate subjects. Documentation regarding research involving the use of human subjects, hazardous materials, vertebrate animals, or endangered species should be included where applicable. (See AAG Chapter VI.B and GPG Chapter II.D.5 and II.D.6).

d. Letters of commitment. Documentation of collaborative arrangements of significance to the proposal through letters of commitment should be included. (See GPG Chapter II.C.2.d.(iv)). All letters of commitment to collaborate must provide specific information regarding the collaboration, including whether it involves sharing resources (data, access to computational resources, or use of other equipment), time and effort, etc. No other type of letter can be provided. The lead institution provides the letters of commitment to collaboration. Proposals that plan to utilize specific resources such as, say, the NSF Cloud platforms, will be expected to provide letters of commitment from the respective organization.

e. Postdoctoral mentoring plan. This one-page supplementary document, describing how postdoctoral researchers will be mentored, is required of all proposals that will have postdoctoral researchers funded by the grant. The lead institution provides this mentoring plan for the entire project. Reviewers will be asked to review the mentoring plan, as appropriate.

Proposals that do not comply with these requirements will be returned without review.

B. Budgetary Information

Cost Sharing: Inclusion of voluntary committed cost sharing is prohibited.

Budget Preparation Instructions:

It is expected that at least one PI from each funded project will attend an annual BIGDATA Principal Investigator (PI) meeting to present project research findings and broader impacts. Collaborative proposals involving investigators from multiple disciplines should have PIs from each of the major disciplines for the grant attending the annual meeting. Requested budgets should include funds for travel to this event.

C. Due Dates

- Full Proposal Deadline(s) (due by 5 p.m. proposer's local time):
  
  May 20, 2015

  Full Proposal Due

D. FastLane/Grants.gov Requirements

For Proposals Submitted Via FastLane:

To prepare and submit a proposal via FastLane, see detailed technical instructions available at: https://www.fastlane.nsf.gov/a1/newstan.htm. For FastLane user support, call the FastLane Help Desk at 1-800-673-6188 or e-mail fastlane@nsf.gov. The FastLane Help Desk answers general technical questions related to the use of the FastLane system. Specific questions related to this program solicitation should be referred to the NSF program staff contact(s) listed in Section VIII of this funding opportunity.

For Proposals Submitted Via Grants.gov:

Before using Grants.gov for the first time, each organization must register to create an institutional profile. Once registered, the applicant's organization can then apply for any federal grant on the Grants.gov website. Comprehensive information about using Grants.gov is available on the Grants.gov Applicant Resources webpage: http://www.grants.gov/web/grants/applicants.html. In addition, the NSF Grants.gov Application Guide (see link in Section V.A) provides instructions regarding the technical preparation of proposals via Grants.gov. For Grants.gov user support, contact the Grants.gov Contact Center at 1-800-518-4726 or by email: support@grants.gov. The Grants.gov Contact Center answers general technical questions related to the use of Grants.gov. Specific questions related to this program solicitation should be referred to the NSF program staff contact(s) listed in Section VIII of this solicitation.

Submitting the Proposal: Once all documents have been completed, the Authorized Organizational Representative (AOR) must submit the application to Grants.gov and verify the desired funding opportunity and agency to which the application is submitted. The AOR must then sign and submit the application to Grants.gov. The completed application will be transferred to the NSF FastLane system for further processing.

Proposers that submitted via FastLane are strongly encouraged to use FastLane to verify the status of their submission to NSF. For proposers that submitted via Grants.gov, until an application has been received and validated by NSF, the Authorized Organizational Representative may check the status of an application on Grants.gov. After proposers have received an e-mail notification from NSF, Research.gov should be used to check the status of an application.

VI. NSF PROPOSAL PROCESSING AND REVIEW PROCEDURES

Proposals received by NSF are assigned to the appropriate NSF program for acknowledgement and, if they meet NSF requirements, for review. All proposals are carefully reviewed by a scientist, engineer, or educator serving as an NSF Program Officer, and usually
by three to ten other persons outside NSF either as ad hoc reviewers, panelists, or both, who are experts in the particular fields represented by the proposal. These reviewers are selected by Program Officers charged with oversight of the review process. Proposers are invited to suggest names of persons they believe are especially well qualified to review the proposal and/or persons they would prefer not review the proposal. These suggestions may serve as one source in the reviewer selection process at the Program Officer’s discretion. Submission of such names, however, is optional. Care is taken to ensure that reviewers have no conflicts of interest with the proposal. In addition, Program Officers may obtain comments from site visits before recommending final action on proposals. Senior NSF staff further review recommendations for awards. A flowchart that depicts the entire NSF proposal and award process (and associated timeline) is included in the GPG as Exhibit III-1.

A comprehensive description of the Foundation's merit review process is available on the NSF website at: http://nsf.gov/bfa/dias/policy/merit_review/.

Proposers should also be aware of core strategies that are essential to the fulfillment of NSF’s mission, as articulated in Empowering the Nation Through Discovery and Innovation: NSF Strategic Plan for Fiscal Years (FY) 2011-2016. These strategies are integrated in the program planning and implementation process, of which proposal review is one part. NSF’s mission is particularly well-implemented through the integration of research and education and broadening participation in NSF programs, projects, and activities.

One of the core strategies in support of NSF’s mission is to foster integration of research and education through the programs, projects and activities it supports at academic and research institutions. These institutions provide abundant opportunities where individuals may concurrently assume responsibilities as researchers, educators, and students, and where all can engage in joint efforts that infuse education with the excitement of discovery and enrich research through the variety of learning perspectives.

Another core strategy in support of NSF’s mission is broadening opportunities and expanding participation of groups, institutions, and geographic regions that are underrepresented in STEM disciplines, which is essential to the health and vitality of science and engineering. NSF is committed to this principle of diversity and deems it central to the programs, projects, and activities it considers and supports.

A. Merit Review Principles and Criteria

The National Science Foundation strives to invest in a robust and diverse portfolio of projects that creates new knowledge and enables breakthroughs in understanding across all areas of science and engineering research and education. To identify which projects to support, NSF relies on a merit review process that incorporates consideration of both the technical aspects of a proposed project and its potential to contribute more broadly to advancing NSF’s mission “to promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national defense; and for other purposes.” NSF makes every effort to conduct a fair, competitive, transparent merit review process for the selection of projects.

1. Merit Review Principles

These principles are to be given due diligence by PIs and organizations when preparing proposals and managing projects, by reviewers when reading and evaluating proposals, and by NSF program staff when determining whether or not to recommend proposals for funding and while overseeing awards. Given that NSF is the primary federal agency charged with nurturing and supporting excellence in basic research and education, the following three principles apply:

- All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.
- NSF projects, in the aggregate, should contribute more broadly to achieving societal goals. These “Broader Impacts” may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. The project activities may be based on previously established and/or innovative methods and approaches, but in either case must be well justified.
- Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects. If the size of the activity is limited, evaluation of that activity in isolation is not likely to be meaningful. Thus, assessing the effectiveness of these activities may best be done at a higher, more aggregated, level than the individual project.

With respect to the third principle, even if assessment of Broader Impacts outcomes for particular projects is done at an aggregated level, PIs are expected to be accountable for carrying out the activities described in the funded project. Thus, individual projects should include clearly stated goals, specific descriptions of the activities that the PI intends to do, and a plan in place to document the outputs of those activities.

These three merit review principles provide the basis for the merit review criteria, as well as a context within which the users of the criteria can better understand their intent.

2. Merit Review Criteria

All NSF proposals are evaluated through use of the two National Science Board approved merit review criteria. In some instances, however, NSF will employ additional criteria as required to highlight the specific objectives of certain programs and activities.

The two merit review criteria are listed below. Both criteria are to be given full consideration during the review and decision-making processes; each criterion is necessary but neither, by itself, is sufficient. Therefore, proposers must fully address both criteria. (GPG Chapter II.C.2.d.i. contains additional information for use by proposers in development of the Project Description section of the proposal.) Reviewers are strongly encouraged to review the criteria, including GPG Chapter II.C.2.d.i., prior to the review of a proposal.

When evaluating NSF proposals, reviewers will be asked to consider what the proposers want to do, why they want to do it, how they plan to do it, how they will know if they succeed, and what benefits could accrue if the project is successful. These issues apply both to the technical aspects of the proposal and the way in which the project may make broader contributions. To that end, reviewers will be asked to evaluate all proposals against two criteria:

- Intellectual Merit: The Intellectual Merit criterion encompasses the potential to advance knowledge; and
- Broader Impacts: The Broader Impacts criterion encompasses the potential to benefit society and contribute to the achievement of specific, desired societal outcomes.

The following elements should be considered in the review for both criteria:

1. What is the potential for the proposed activity to
B. Award Conditions

Investigator. (See Section VI.B for additional information on the review process).

Verbatim copies of reviews, not including the identity of the reviewer, will be provided automatically to the Principal Investigator. Organizations whose proposals are declined will be advised as promptly as possible by the cognizant NSF Program administering the program. Proposers are reminded that only a Grants and Agreements Officer may make commitments, obligations or awards on behalf of NSF or authorize the expenditure of funds. No commitment on the part of NSF should be inferred from the deadline or target date, or receipt date, whichever is later. The interval ends when the Division Director accepts the Program Officer's recommendation.

For this solicitation, broader impacts may include infrastructure development appropriate to advancing the development of novel techniques and technologies. Infrastructure support may: transform interdisciplinary and collaborative research, enable interdisciplinary training, enable long term stewardship of data and software; provide for sustainable shared infrastructure and ensure the availability of data, source code and software to the scientific community and general public.

Additional Solicitation Specific Review Criteria

1. Additional Intellectual Merit Criteria

BIGDATA proposals should be well-motivated via application use cases and/or characteristics of real datasets, in the case of Foundational proposals; and, should include true partnerships with domain science and engineering, in the case of Innovative Applications proposals. Proposals will be evaluated based on how well they match these criteria.

2. Additional Broader Impact Criteria

For this solicitation, broader impacts may include infrastructure development appropriate to advancing the development of novel techniques and technologies.

B. Review and Selection Process

Proposals submitted in response to this program solicitation will be reviewed by Ad hoc Review and/or Panel Review.

NSF will coordinate and manage the review of proposals. Relevant information about proposals and reviews of proposals will be shared with OFR program staff as appropriate. Proposals selected for support by OFR will be awarded by NSF using funds transferred from OFR.

After scientific, technical and programmatic review and consideration of appropriate factors, the NSF Program Officer recommends to the cognizant Division Director whether the proposal should be declined or recommended for award. NSF is striving to be able to tell applicants whether their proposals have been declined or recommended for funding within six months. The time interval begins on the deadline or target date, or receipt date, whichever is later. The interval ends when the Division Director accepts the Program Officer's recommendation.

A summary rating and accompanying narrative will be completed and submitted by each reviewer. In all cases, reviews are treated as confidential documents. Verbatim copies of reviews, excluding the names of the reviewers, are sent to the Principal Investigator/Project Director by the Program Officer. In addition, the proposer will receive an explanation of the decision to award or decline funding.

In all cases, after programmatic approval has been obtained, the proposals recommended for funding will be forwarded to the Division of Grants and Agreements for review of business, financial, and policy implications and the processing and issuance of a grant or other agreement. Proposers are cautioned that only a Grants and Agreements Officer may make commitments, obligations or awards on behalf of NSF or authorize the expenditure of funds. No commitment on the part of NSF should be inferred from technical or budgetary discussions with a NSF Program Officer. A Principal Investigator or organization that makes financial or personnel commitments in the absence of a grant or cooperative agreement signed by the NSF Grants and Agreements Officer does so at their own risk.

VII. AWARD ADMINISTRATION INFORMATION

A. Notification of the Award

Notification of the award is made to the submitting organization by a Grants Officer in the Division of Grants and Agreements. Organizations whose proposals are declined will be advised as promptly as possible by the cognizant NSF Program administering the program. Verbatim copies of reviews, not including the identity of the reviewer, will be provided automatically to the Principal Investigator. (See Section VI.B. for additional information on the review process).

B. Award Conditions
An NSF award consists of: (1) the award notice, which includes any special provisions applicable to the award and any numbered amendments thereto; (2) the budget, which indicates the amounts, by categories of expense, on which NSF has based its support (or otherwise communicates any specific approvals or disapprovals of proposed expenditures); (3) the proposal referenced in the award notice; (4) the applicable award conditions, such as Grant General Conditions (GC-1)*; or Research Terms and Conditions* and (5) any announcement or other NSF issuance that may be incorporated by reference in the award notice. Cooperative agreements also are administered in accordance with NSF Cooperative Agreement Financial and Administrative Terms and Conditions (CA-FATC) and the applicable Programmatic Terms and Conditions. NSF awards are electronically signed by an NSF Grants and Agreements Officer and transmitted electronically to the organization via e-mail.

*These documents may be accessed electronically on NSF's Website at http://www.nsf.gov/awards/managing/award_conditions.jsp?org=NSF. Paper copies may be obtained from the NSF Publications Clearinghouse, telephone (703) 292-7827 or by e-mail from nsfpubs@nsf.gov.


C. Reporting Requirements

For all multi-year grants (including both standard and continuing grants), the Principal Investigator must submit an annual project report to the cognizant Program Officer at least 90 days prior to the end of the current budget period. (Some programs or awards require submission of more frequent project reports). Within 90 days following expiration of a grant, the PI also is required to submit a final project report, and a project outcomes report for the general public.

Failure to provide the required annual or final project reports, or the project outcomes report, will delay NSF review and processing of any future funding increments as well as any pending proposals for all identified PIs and co-PIs on a given award. PIs should examine the formats of the required reports in advance to assure availability of required data.

Pis are required to use NSF’s electronic project-reporting system, available through Research.gov, for preparation and submission of annual and final project reports. Such reports provide information on accomplishments, project participants (individual and organizational), publications, and other specific products and impacts of the project. Submission of the report via Research.gov constitutes certification by the PI that the contents of the report are accurate and complete. The project outcomes report also must be prepared and submitted using Research.gov. This report serves as a brief summary, prepared specifically for the public, of the nature and outcomes of the project. This report will be posted on the NSF website exactly as it is submitted by the PI.


VIII. AGENCY CONTACTS

Please note that the program contact information is current at the time of publishing. See program website for any updates to the points of contact.

General inquiries regarding this program should be made to:

- Chaitanya Baru, Senior Advisor- CISE/OAD, telephone: (703) 292-4541, email: cbaru@nsf.gov
- Sylvia Spengler, Program Director - CISE/IIS, telephone: (703) 292-8930, email: sspengle@nsf.gov
- Jack Snoeyink, Program Director - CISE/CCF, 1115, telephone: (703) 292-8910, email: jsnoeyin@nsf.gov
- Amy Apon, Program Director - CISE/CNS, telephone: 703-292-8950, email: aapon@nsf.gov
- Elizabeth R. Blood, Program Director - BIO/OAD, telephone: (703) 292-4349, email: eblood@nsf.gov
- Helen T. Martin, Program Director- EHR/DRL, telephone: (703) 292-7016, email: htmartin@nsf.gov
- George Haddad, Program Director - ENG/ECCS, telephone: (703) 292-8897, email: ghaddad@nsf.gov
- Mona Zaghoul, Program Director - NEG/ECCS, telephone: (703) 292-8339, email: mzaghoul@nsf.gov
- Eva Zanzkeria, Program Director - GEO/OAD, telephone: (703) 292-4734, email: ezanzker@nsf.gov
- Xiaoming Huo, Program Director - MPS/DMS, telephone: (703) 292-4876, email: xhuo@nsf.gov
- Nandini Kannan, Program Director - MPS/DMS, telephone: (703) 292-8104, email: nakannan@nsf.gov
- Bogdan Mihaila, Program Director - MPS/PHY, telephone: (703) 292-8235, email: bmihaila@nsf.gov
- Heng Xu, Program Director - SBE/SES, telephone: (703) 292-8643, email: hxu@nsf.gov
- Edward Garnett, Program Manager -- OFR/Dept. of Treasury, telephone: (202) 927-8025, email: edward.garnett@treasury.gov

For questions related to the use of FastLane, contact:

- FastLane Help Desk, telephone: 1-800-673-6188; e-mail: fastlane@nsf.gov.

For questions relating to Grants.gov contact:

- Grants.gov Contact Center: If the Authorized Organizational Representatives (AOR) has not received a confirmation message from Grants.gov within 48 hours of submission of application, please contact via telephone: 1-800-518-4726; e-
IX. OTHER INFORMATION

The NSF website provides the most comprehensive source of information on NSF Directorates (including contact information), programs and funding opportunities. Use of this website by potential proposers is strongly encouraged. In addition, "NSF Update" is an information-delivery system designed to keep potential proposers and other interested parties apprised of new NSF funding opportunities and publications, important changes in proposal and award policies and procedures, and upcoming NSF Grants Conferences. Subscribers are informed through e-mail or the user's Web browser each time new publications are issued that match their identified interests. "NSF Update" also is available on NSF's website at https://public.govdelivery.com/accounts/USNSF/subscriber/new?topic_id=USNSF_179.

Grants.gov provides an additional electronic capability to search for Federal government-wide grant opportunities. NSF funding opportunities may be accessed via this mechanism. Further information on Grants.gov may be obtained at http://www.grants.gov.

ABOUT THE NATIONAL SCIENCE FOUNDATION

The National Science Foundation (NSF) is an independent Federal agency created by the National Science Foundation Act of 1950, as amended (42 USC 1861-75). The Act states the purpose of the NSF is "to promote the progress of science; [and] to advance the national health, prosperity, and welfare by supporting research and education in all fields of science and engineering."

NSF funds research and education in most fields of science and engineering. It does this through grants and cooperative agreements to more than 2,000 colleges, universities, K-12 school systems, businesses, informal science organizations and other research organizations throughout the US. The Foundation accounts for about one-fourth of Federal support to academic institutions for basic research.

NSF receives approximately 55,000 proposals each year for research, education and training projects, of which approximately 11,000 are funded. In addition, the Foundation receives several thousand applications for graduate and postdoctoral fellowships. The agency operates no laboratories itself but does support National Research Centers, user facilities, certain oceanographic vessels and Arctic and Antarctic research stations. The Foundation also supports cooperative research between universities and industry, US participation in international scientific and engineering efforts, and educational activities at every academic level.

Facilitation Awards for Scientists and Engineers with Disabilities provide funding for special assistance or equipment to enable persons with disabilities to work on NSF-supported projects. See Grant Proposal Guide Chapter II, Section D.2 for instructions regarding preparation of these types of proposals.

The National Science Foundation has Telephonic Device for the Deaf (TDD) and Federal Information Relay Service (FIRS) capabilities that enable individuals with hearing impairments to communicate with the Foundation about NSF programs, employment or general information. TDD may be accessed at (703) 292-5090 and (800) 281-8749, FIRS at (800) 877-8339.

The National Science Foundation Information Center may be reached at (703) 292-5111.

The National Science Foundation promotes and advances scientific progress in the United States by competitively awarding grants and cooperative agreements for research and education in the sciences, mathematics, and engineering.

To get the latest information about program deadlines, to download copies of NSF publications, and to access abstracts of awards, visit the NSF Website at http://www.nsf.gov

- Location: 4201 Wilson Blvd. Arlington, VA 22230
- For General Information (NSF Information Center): (703) 292-5111
- TDD (for the hearing-impaired): (703) 292-5090
- To Order Publications or Forms:
  - Send an e-mail to: nsfpubs@nsf.gov
  - or telephone: (703) 292-7827
- To Locate NSF Employees: (703) 292-5111

PRIVACY ACT AND PUBLIC BURDEN STATEMENTS

The information requested on proposal forms and project reports is solicited under the authority of the National Science Foundation
Act of 1950, as amended. The information on proposal forms will be used in connection with the selection of qualified proposals; and project reports submitted by awardees will be used for program evaluation and reporting within the Executive Branch and to Congress. The information requested may be disclosed to qualified reviewers and staff assistants as part of the proposal review process; to proposer institutions/grantees to provide or obtain data regarding the proposal review process, award decisions, or the administration of awards; to government contractors, experts, volunteers and researchers and educators as necessary to complete assigned work; to other government agencies or other entities needing information regarding applicants or nominees as part of a joint application review process, or in order to coordinate programs or policy; and to another Federal agency, court, or party in a court or Federal administrative proceeding if the government is a party. Information about Principal Investigators may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records," 69 Federal Register 26410 (May 12, 2004), and NSF-51, "Reviewer/Proposal File and Associated Records," 69 Federal Register 26410 (May 12, 2004). Submission of the information is voluntary. Failure to provide full and complete information, however, may reduce the possibility of receiving an award.

An agency may not conduct or sponsor, and a person is not required to respond to, an information collection unless it displays a valid Office of Management and Budget (OMB) control number. The OMB control number for this collection is 3145-0058. Public reporting burden for this collection of information is estimated to average 120 hours per response, including the time for reviewing instructions. Send comments regarding the burden estimate and any other aspect of this collection of information, including suggestions for reducing this burden, to:

Suzanne H. Plimpton
Reports Clearance Officer
Office of the General Counsel
National Science Foundation
Arlington, VA 22230