# Next Generation Cyberinfrastructure Tools
## With Applications to Complex Behavior of Organizations and Individuals

**Program Solicitation**
**NSF 05-563**

**National Science Foundation**
Directorate for Social, Behavioral, and Economic Sciences
    Division of Behavioral and Cognitive Sciences
    Division of Social and Economic Sciences
Directorate for Computer and Information Science and Engineering
Office of Cyberinfrastructure

**Full Proposal Deadline(s)** (due by 5 p.m. submitter's local time):

May 30, 2005

## SUMMARY OF PROGRAM REQUIREMENTS

## General Information

**Program Title:**

Next Generation Cyberinfrastructure Tools
With Applications to Complex Behavior of Organizations and Individuals

**Synopsis of Program:**

Researchers in the social and behavioral sciences and computer and information sciences have many important synergistic relationships.One way in which this is manifest is in the development and utilization of data. On the one hand, social and behavioral scientists find new ways to create and analyze data in their endeavors to describe human and organizational behavior. On the other hand, computer and information scientists conduct research that yields new ways to improve both domain-specific and general-purpose tools to analyze and visualize scientific data -- such as improving processing power, enhanced interoperability of data from different sources, data mining, data integration, information indexing and data confidentiality protection - or what we have termed cyberinfrastructure tools.

This solicitation invites proposals for **"information infrastructure testbeds"**, each of which would include the development of the next generation of cyberinfrastructure tools applied to data from various sources collected in two areas of research fundamental to social and behavioral scientists: **organizations** and **individuals.** The tools that are developed on these platforms must not only change ways in which social and behavioral scientists research the behavior of organizations and individuals, but also serve sciences more broadly.

It is envisioned that proposals for the "**organization information testbed"** will address three specific components:

- the development of tools that facilitate the integration of qualitative and quantitative information from heterogeneous sources, multiple media, and/or multiple modes;
- investment in basic research that addresses the protection of the confidentiality of respondents in computerized, widely accessible databases; and
- the development of incentives, standards and policies for collecting, storing, archiving, accessing, and publishing research results using organization-relevant information.

It is envisioned that proposals for the **"individual information testbed"** should concern cyberinfrastructure tools that can be applied to both large scale and distributed data-sets. Proposals should address cyberinfrastructure tools that facilitate automatic collection, integration, annotation, archiving, accessing, and analyzing of

- existing distributed data sets and/or
- extensive audio and video recordings and details of physical artifacts, while paying special attention to
- the protection of the confidentiality of participant identity in widely accessible, computerized databases.

**Cognizant Program Officer(s):**

- Miriam Heller, Program Director, Office of the Director, Office of Cyberinfrastructure, 1145 S, telephone: (703) 292-7025, fax: (703) 292-9060, email: mheller@nsf.gov
- Lawrence E. Brandt, Program Manager, Directorate for Computer & Information Science & Engineering, Division of Information and Intelligent Systems, 1125 S, telephone: (703) 292-8930, fax: (703) 292-9073, email: lbrandt@nsf.gov
- Joan Maling, Program Director, Directorate for Social, Behavioral & Economic Sciences, Division of Behavioral and Cognitive Sciences, 995 N, telephone: (703) 292-8046, fax: (703) 292-9068, email: jmaling@nsf.gov
- Jacqueline Meszaros, Program Director, Directorate for Social, Behavioral & Economic Sciences, Division of Social and Economic Sciences, 995 N, telephone: (703) 292-7261, fax: (703) 292-9068, email: jmeszaro@nsf.gov
- Kevin L. Thompson, Program Director, Office of the Director, Office of Cyberinfrastructure, 1145 S, telephone: (703) 292-8962, fax: (703) 292-9060, email: kthompso@nsf.gov

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

- 47.070 --- Computer and Information Science and Engineering
- 47.080 --- Office of Cyberinfrastructure
- 47.075 --- Social, Behavioral and Economic Sciences

## Eligibility Information

- **Organization Limit:** Universities or colleges, including two- and four-year colleges and community colleges, acting on behalf of their faculty members may submit proposals. In addition, non-profit non-academic organizations, such as independent museums, observatories, research laboratories, professional societies and similar organizations in the US that are directly associated with educational or research activities, may submit proposals. NSF encourages proposals for collaboration with international researchers, for-profit corporations, and national laboratories. For-profit organizations, government laboratories, and foreign organizations may not apply directly; however, they may participate in subawards. Such subawards should be justified by explaining the unique capabilities being made available.

- **PI Eligibility Limit:** None Specified.
- **Limit on Number of Proposals:** None Specified.

## Award Information

- **Anticipated Type of Award:** Standard or Continuing Grant
- **Estimated Number of Awards:** 2 (One for each testbed, approximating $2 million each).
- **Anticipated Funding Amount:** $4,000,000 pending the availability of funds

## Proposal Preparation and Submission Instructions

### A. Proposal Preparation Instructions

- **Full Proposal Preparation Instructions:** This solicitation contains information that deviates from the standard Grant Proposal Guide (GPG) proposal preparation guidelines. Please see the full text of this solicitation for further information.

### B. Budgetary Information

- **Cost Sharing Requirements:** Cost Sharing is not required by NSF.
- **Indirect Cost (F&A) Limitations:** Not Applicable.
- **Other Budgetary Limitations:** Not Applicable.

## C. Due Dates

- **Full Proposal Deadline Date(s)** (due by 5 p.m. submitter's local time):
  May 30, 2005

## Proposal Review Information

- **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

Breakthroughs in the social and behavioral sciences have been enabled by and increasingly depend on advances in cybertechnologies. The impact of social and behavioral research has proven to be substantial – leading, for example, to new insights into core national challenges such as the creation of wealth and jobs, as well as to fundamental understanding of

3

behavioral phenomena at scales ranging from the molecular to the social. The special technical needs of the social and behavioral scientists have also posed numerous research challenges to cyberinfrastucture experts and contributed to substantial advances in the cybertechnologies of other sciences: for example, geographic information systems applications used in engineering, the geosciences and biology, and statistical software packages developed for social and behavioral science applications but now widely used across the health sciences. Yet, the range of opportunities and the full benefits that can be derived from the application and use of large-scale scientific infrastructure, currently called "cyberinfrastructure" at NSF, have yet to be fully explored and realized by social and behavioral scientists.

Stimulated by the potential contribution of the social and behavioral sciences to such national challenges, as well as by the potential advances that can be generated by responding to the special needs of social and behavioral scientists, the Directorate for Social, Behavioral and Economic Sciences (SBE) together with the Directorate for Computer and Information Science and Engineering (CISE) announces a competition designed to develop new cyberinfrastructure that can be shared across the social and behavioral sciences as well as with other sciences. Research will be supported within two "testbeds", or platforms on which experimental cyberinfrastructure tools and products will be deployed with the aim of identifying successful tools and products for further development. Examples of such cyberinfrastructure tools might be shared domain-specific and general-purpose tools to analyze and visualize scientific data -- such as improving processing power, the interoperability of data from different sources, data mining, data integration, information indexing and data confidentiality protection.

These "testbeds" are:

I. information collected on **organizations** from a variety of heterogeneous, independently developed data sources, such as administrative and survey data, temporal, spatial and image data or textual data. The goal is to free users from having to locate the data sources, interact with each data source in isolation, and manually combine data from multiple formats and multiple sources. This could be achieved through the creation of new and more accurate and efficient ways to collect, code and analyze qualitative information from case studies, and other sources, and to enable the linking of this information with repositories of quantitative data, while protecting fundamental privacy and confidentiality concerns. The research should be designed to show how appropriate cyberinfrastructure tools can lead to multiple advances in the empirical understanding of how organizations emerge, develop, thrive or weaken.

II information collected on **individuals** such as human cognition, activities and artifacts. The focus will be on projects to develop cyberinfrastructure tools that can be used to better acquire, annotate, archive, access and analyze data in the form of audio and video recordings, records of physical artifacts, changes in images across time, etc., and that boost levels of volume and complexity tractable for analysis on both large-scale and distributed datasets. These proposals should show how appropriate cyberinfrastructure tools can lead to multiple advances in the detailed empirical understanding of individual behavior and/or the integration of previously isolated distributed databases.

Advances in both "testbeds" must have applicability across and beyond the various social and behavioral sciences. They must also have the demonstrated potential to be scalable across time, geographies, institutions, and/or disciplines. Finally, they must exploit state-of-the-art knowledge to ensure sustained relevance and usability.

Additional information on cyberinfrastructure and the social, behavioral, and economic sciences can be found in "Revolutionizing Science and Engineering through Cyberinfrastructure: Report on the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastruture, January 2003, Daniel E. Atkins, Chair"

http://www.nsf.gov/cise/sci/reports/toc.jsp

and the "ACLS Commission on Cyberinfrastructure for the Humanities & Social Sciences,"

http://www.acls.org/cyberinfrastructure/cyber_meeting_notes_august.htm

## II. PROGRAM DESCRIPTION

This competition is viewed as an initial step toward meeting the goal of innovatively developing cyberinfrastructure tools that are applied to two "testbeds" and yet have broad applicability across the social, behavioral, economic and other sciences and engineering. *Investigators are strongly encouraged to contact the cognizant program officers.*

Proposals must address the protection of data providers from identification, exploitation, and other misuses of personal or organizational information. Such misuses present a perpetual challenge to the melding of data and media of different types in a tool for widespread use. Proposals in response to this solicitation must show a sophisticated understanding of this

sociotechnical problem and must propose to advance fundamental knowledge of effective privacy protections during the development of the analytical tools and in their later use by various research communities.

Proposals must demonstrate potential long-term sustainability, usability, and impact. This could be achieved for the organizational "testbed", for example, by documenting proposed collaboration with firms in an industry, attracting support from foundations or developing replicable incentive-compatible policies for collecting, storing, accessing, and disseminating data while continuing to utilize and advance relevant cybertechnology. It could be achieved for both "testbeds" by, for example, capitalizing on existing infrastructure, such as research centers or supercomputer centers.

Research supported under this solicitation must involve medium-scale groups of senior personnel representing multiple disciplines, including social and behavioral science as well as computer science. Teams must include individuals with a demonstrated capacity to identify and develop the appropriate cyberinfrastructure tools -- which is what we term methodologies and tools for the representation and manipulation of large volumes of data for heterogeneous, multimodal sources, on either organizations or individuals. All proposals should indicate how the research will provide opportunities for training, education and broadening participation.

Examples of topics the cyberinfrastructure tools could be applied to are listed below; followed by a description of the "testbeds" to which the tools should be applied. It is desirable that tools within a "testbed" are well integrated with each other.

**Examples of Cyberinfrastructure Tools**

A. Social and Behavioral Informatics:

- *Social and Behavioral Data Models and System*s. Implementable theoretical foundations for the representation and manipulation of advanced data types (e.g., temporal, spatial and image data, textual data, administrative or survey data); data/knowledge calibration and validation. Systems issues include system extensibility; development of user-transparent, multi-level storage management; multi-media data indexing; partial match retrieval algorithms; archiving; and version control. Research in this area must consider the special data and information characteristics associated with the social and behavioral sciences "testbeds" described below.
- *Analysis of Organizational and Individual Databases and Information Resources.* Topics span computing environment transparency; establishing baseline patterns, data examination, selection, analysis and manipulation of temporally or spatially related data; knowledge discovery algorithms; information extraction (e.g., from abstracts of publications), citation analysis, visualization; parallel model execution and cross-validation on large volumes of data; automated knowledge acquisition; incorporation of new knowledge into a system; and audit trail provisions including data provenance. The research in this area must be done in connection with the "testbeds" described below.
- *Analysis of Social and Behavioral Multimodal and Multimedia*. A key research challenge in many research problems is to derive measurements or abstract features from 2-D, 3-D and multispectral images as well as well as other audio and other sensory data and to use this derived information for generating or evaluating hypotheses.
- *Shared Resources Environments*. The construction of shared, archived, and documented data, publication, or software resources that can accelerate the rate of scientific discovery.

B. Information Integration:

The goal of information integration should be to integrate many different, disparate and possibly distributed sources; support automated discovery of new data sources and information within them; facilitate their configuration, management and system maintenance; incorporate structured, semi-structured, text, image, video, time-series, 3D images, citations, graphs, speech and other data streams; and provide flexible querying of the sources and the data.

Some of the specific challenges include:

- *Unifying Data Models and System Descriptions*: There is a need to develop stronger theoretical foundations for the representation and integration of information of various types from extant data models (e.g., temporal, spatial and image data, textual data, administrative and survey data) as well as the scientific literature into conceptually coherent views. Specific topics include: metadata management and integration; the automated collection of metadata from instruments and processes that transform data, ontologies and taxonomies; data/knowledge calibration; heterogeneity of data type and format; scale of distributed systems; rapid integration of new information sources. Research in this area must consider the special data characteristics associated with the social and behavioral sciences disciplines and data described in the "testbeds" below.
- *Reconciling heterogeneous formats schemas and ontologies*: The fundamental problem in any data sharing application is that systems are heterogeneous in many different aspects, such as different ways of representing data and/or knowledge about the world, different representation mechanisms (e.g., relational databases, legacy systems, XML schemas, ontologies), different access methods and policies. In order to share data among heterogeneous sources, approaches to form a semantic mapping of their respective representations are needed to avoid manual

intervention in each step of converting and merging data resources.

- *Web semantics*: Data on the web needs to be defined and linked in a way that it can be used by machines not just for display purposes, but also for automation, integration and reuse of data across various applications. Supported research topics will include frameworks for describing resources, methods of automating inferences about web data and resources, and the development of interoperable ontologies, mark up languages and representations for specific social, behavioral and other scientific domains.
- *Decentralized data-sharing:* Traditional data integration systems use a centralized mediation approach, in which a centralized mediator, employing a mediated schema, accepts user queries and reformulates them over the schemas of the different sources. However, mediated schemas are often hard to agree upon, construct and maintain. For example, researchers conducting social and behavioral research share their experimental results with each other, but may do it in an ad hoc fashion. A similar scenario is found in data sharing among government agencies. Architectures and protocols that enable large-scale sharing of data with no central control are needed.
- *Data-sharing on advanced cyberinfrastructure*: Research topics will include models for federating information resources in advanced grid computing and/or Web services, integration and understanding of sensor information, the collection of metadata from sensors including models and tools to cope with the scale, pervasiveness, concurrency and redundancy of sensor data. Effective integration of network management information will be critical to enable basic networking functions such as routing, overlay node placement, denial-of-service detection, and fault recovery. The integration of network management information will facilitate adapting network resources to changing conditions.
- *On-the-fly integration*: Currently, data integration systems rely on relatively static configurations with a set of long-lived data sources. On-the-fly integration refers to scenarios where one wants to integrate data from a source immediately after discovering it. The challenge is to significantly reduce the time and skill needed to integrate data sources so that scientists can focus on domain problems instead of information technology challenges.

Separate criteria for each "testbed" follow.

## "TESTBED" I: ORGANIZATIONS

Research on organizations has historically involved many different paths using different types of methods and data. One research path has relied upon the collection and analysis of large scale, nationally representative datasets on organizations. Another research path has relied upon the development of rich case study information by individual researchers and sets of researchers. Another has been the development of complex models with rich descriptive realism as exemplified by agent-based modelling approaches; and yet another the examination of information collected by observational and ethnographic methods.

Each of these approaches has led to interesting and useful findings. Researchers working with large-scale firm-level datasets, for example, have uncovered important underlying dynamics: much of productivity growth is due to factor reallocation from less productive to more productive firms; a disproportionate amount of economic change is driven by new firms displacing old; and aggregate economic statistics do not capture the enormous amount of turbulence uncovered by an examination of firm level micro-data. However, the quantitative data are not rich on non-economic indicators, such as intra-organizational networks, personnel policies and the status of women and minorities. In contrast, case study researchers have described rich contextual variation in organizational decision-making, as well as how hard-to-quantify phenomena like culture and management affect organizational performance. Case studies have often uncovered important forces that are later measured in larger-scale work. But case study research has often suffered from the criticism that that work is neither generalizable nor replicable. Similarly, researchers in the agency-based modelling literature who use real-world databases to both calibrate and validate theoretical models as well as generate new data, provide fresh new perspectives on describing organizations - particularly the medium-term dynamics. Yet they need new approaches to organize and analyze data from a variety of sources in order that their models more closely approximate reality . All sets of researchers could gain from the development of cyberinfrastructure tools that permitted the integration of the different sources of data.

Advances in cyber-infrastructure now make it possible to fundamentally change the way in which social scientists collect, create, and analyze data on organizations and people so that each research path can gain from knowledge derived from others. This competition aims to create tools that facilitate the combination of information from a variety of sources -- such as text, documents, video, audio and maps -- using data on organizations as a test-bed. This will involve creating a set of tools that permit the creation, standardization and searchability of a data archive – although the specific context in this proposal is to inform and enrich the empirical analysis of organizations.

Proposed projects should combine elements of each of the following three areas, although consideration will be given to outstanding proposals that address one or two of the areas.

- The development of tools that combine information on organizations from multiple media – both qualitative and quantitative
- Basic research that addresses the protection of the confidentiality of individual and organizational data providers and survey respondents, and
- The development of incentives, standards and policies for collecting, storing, accessing, and publishing research results using the data and information, as well as including appropriate protections of confidentiality and security.

Proposals should focus specifically on the development of tools and methodology for one set of organizations, such as an industry, or group of industries, but the tools developed in the proposed pilot project should be demonstrably scalable to additional industries and organizations (such as unions, voluntary organizations schools, governments and political groups). The tools should have the potential to meet needs in other sciences. The proposed project should document the potential for sustainability. Expressions of interest in the research and the "testbed" area from the private sectors such as, for example, letters from private foundations or potential industrial partners is one way to document such potential. The case for sustainability might also be made by documenting how the proposed tools could be used to develop a database that addresses some of the following research questions or questions of similar national or scientific importance:

- The factors and processes that contribute to the emergence and growth of businesses and successful entrepreneurship - particularly job and wealth creation
- The factors and processes that contribute to the success or failure of mergers and acquisitions
- The organizational response to globalization, such as outsourcing, and the consequences for job creation and business competitiveness
- The factors and processes that contribute to employees' satisfaction or dissatisfaction with their jobs
- The existence and function of internal and boundary spanning networks
- The implications of legal and administrative regulation for customer or employee relationship as well as for the pace or shape of firm development
- The hiring, training and retention of different types of workers: for example, older or disabled individuals

**"TESTBED" II: INDIVIDUALS**.

Research on human behavior has traditionally involved summary data that reduce the complexity of detail available for analysis and subsequent reanalysis, such as summary data that omits details of change across time or variability in individual responses. Additionally, other data collected on small scales would have greater scientific value if used in larger scale studies that combined many data sets. This competition encourages the development of cyberinfrastructure tools that make such data comprehensible, while protecting privacy and confidentiality. Proposals are invited that address the complex behavior of individuals. For example, proposals that integrate multiple data types, such as voice, video, images, or physical artifacts, and that acquire, annotate, archive, access and analyze these and boost levels of volume and complexity up to and including large-scale and distributed datasets.

For example, real time observations of human behavior can generate vast data bases that currently exist in multiple locations on and off the web. Or consider legislation that mandates anthropological mitigation during construction projects, resulting in the excavation of millions of individual objects that lack integration into a single data-base. A similar story can be told for data on children's development and social phenomena as well as remote sensing images that inform work in geography and the regional sciences. In addition there are huge numbers of language data sets and knowledge bases, variously sized, that are currently available on the Web or that will soon come online.

As another example, archaeological artifacts, fossils and skeletal material, and historical records all provide threads of information about past human actions and behavior. The human migrations that populated the world, the rise and fall of civilizations, the origin and spread of innovative behaviors such as agriculture are all represented by data from which many insights could be gained if standardization and interoperability could be applied to extant data collections. In many cases a means to incorporate data and analytical techniques from a wide variety of other disciplines (e.g. geographic, geological, genetic, linguistic) would greatly enhance the utility of all of the constituent parts. The ability to incorporate ethnographic case material would enrich large, cross-sectional databases and add substance and validity.

These "testbed" descriptions are applicable across many disciplines in the behavioral, cognitive, anthropological and geographic sciences, which have amassed large numbers and varieties of data sets and analyses of them. The key to solving the interoperability problem is to develop concepts, semantics and mechanisms that can be used to store and analyze data. Using the data accumulated by linguists as an example, it is clear that these resources are currently in a mix of formats and states of annotation including (a) raw or partially analyzed text files in a variety of formats, (b) idiosyncratically prepared and analyzed dictionaries, word lists, sketch grammars, cross-linguistic typological comparisons,in a variety of formats, flat or relational databases, and other files, (c) for example, collections of annotated text and lexicons in standard "best practice" XML formats, and (d) audio and audio-video recordings of speech, ideally but not necessarily aligned with one or more transcriptions (phonetic, standard orthography), and possibly also with sociolinguistic and/or grammatical annotations.

Data abound, but, even assuming that these resources are discoverable, they are still not interoperable, i.e. comparable for purposes of systematically carrying out research across them, or for enriching them. Achieving interoperability will require that at least the following three components be developed, in addition to the development of procedures and protocols to protect confidentiality and privacy as well as to prevent copyright violations:

- Standard methods for converting legacy and other non-best practice data collections to best practice.

- A universally accessible knowledge base that provides uniform and agreed upon rules for the integrating and interpreting of the data, both domain-specific (e.g. linguistic, social psychology, or geography, etc.) or across domains.
- A set of methods for linking the best-practice formatted data to the standard knowledge base. These involve the creation of term sets and standard protocols for linking them both to the documents that use them and to the general knowledge base.

Parallel issues emerge across many disciplines in the behavioral, cognitive, anthropological and geographic sciences, as well as other sciences. There exist now unlinked data resources requiring the terabyte scale of storage, and analyses from around the world, the terra scale, that, with attention to interoperability, will answer longstanding questions about the complexity of human behavior and generate increasingly complex questions. The use of data that are in the form of visual images, audio and visual records, and physical objects is essential, but not unique to the behavioral cognitive sciences. Thus these fields offer an excellent testbed for the development of cyberinfrastructure for amassing and analyzing this sort of data, but the developing infrastructure will have uses across the sciences broadly. These tools necessarily raise ethical issues including questions of confidentiality, privacy, and long term access that must be addressed in any proposal.

## III. ELIGIBILITY INFORMATION

The categories of proposers identified in the Grant Proposal Guide are eligible to submit proposals under this program announcement/solicitation.

## IV. AWARD INFORMATION

The National Science Foundation expects to make one standard or continuing grant in each "testbed", of approximately $2 million each, for a total of $4 million, to cover these activities, subject to the availability of funds. NSF support may be requested for professional staff since such participation is likely to be necessary.

## V. PROPOSAL PREPARATION AND SUBMISSION INSTRUCTIONS

## A. Proposal Preparation Instructions

**Full Proposal Instructions:**

Proposals submitted in response to this program announcement/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 pubs@nsf.gov.

The following instructions supplement the GPG guidelines.

- The project summary must identify whether the proposer is submitting for "testbed" I or "testbed" II.
- The project description may be up to 15 pages. Of these, up to 5 pages should be spent presenting an outreach and dissemination plan, a coordination plan (described below) and an indication that the research will have some teaching and education component.
- The proposal must contain a coordination plan, which includes 1) the specific roles of the PI, co-PIs, other senior personnel and paid consultants at all institutions 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 integration (e.g., yearly workshops, graduate student exchange, project meetings at conferences, use of the grid for videoconferences, software repositories, etc.), and 4) pointers to the budget line items that support these coordination mechanisms. If budget cuts are necessary, NSF staff will make every effort not to reduce the budget for project coordination other than proportionally.
- Proposals may refer to supporting materials, such as the datasets or methodology that may be used, previous related work and other directly relevant information, posted on investigators' publicly available websites. Reviewers are not, however, obligated to view supporting materials.

- Project teams must consist of medium-scale groups led by senior personnel from multiple disciplines, including social and behavioral science as well as computer science.

Proposers are reminded to identify the program announcement/solicitation number (05-563) in the program announcement/solicitation block on the proposal Cover Sheet. Compliance with this requirement is critical to determining the relevant proposal processing guidelines. Failure to submit this information may delay processing.

## B. Budgetary Information

**Cost Sharing:**

Cost sharing is not required by NSF in proposals submitted under this Program Solicitation.

## C. Due Dates

Proposals must be submitted by the following date(s):

**Full Proposal Deadline(s)** (due by 5 p.m. submitter's local time):

May 30, 2005

## D. FastLane Requirements

Proposers are required to prepare and submit all proposals for this announcement/solicitation through the FastLane system. Detailed instructions for proposal preparation and submission via FastLane are 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 announcement/solicitation should be referred to the NSF program staff contact(s) listed in Section VIII of this announcement/solicitation.

*Submission of Electronically Signed Cover Sheets.* The Authorized Organizational Representative (AOR) must electronically sign the proposal Cover Sheet to submit the required proposal certifications (see Chapter II, Section C of the Grant Proposal Guide for a listing of the certifications). The AOR must provide the required electronic certifications within five working days following the electronic submission of the proposal. Proposers are no longer required to provide a paper copy of the signed Proposal Cover Sheet to NSF. Further instructions regarding this process are available on the FastLane Website at: http://www.fastlane.nsf.gov

## VI. PROPOSAL REVIEW INFORMATION

## A. NSF Proposal Review Process

Reviews of proposals submitted to NSF are solicited from peers with expertise in the substantive area of the proposed research or education project. These reviewers are selected by Program Officers charged with the oversight of the review process. NSF invites the proposer to suggest, at the time of submission, the names of appropriate or inappropriate reviewers. Care is taken to ensure that reviewers have no conflicts with the proposer. Special efforts are made to recruit reviewers from non-academic institutions, minority-serving institutions, or adjacent disciplines to that principally addressed in the proposal.

The National Science Board approved revised criteria for evaluating proposals at its meeting on March 28, 1997 (NSB 97-72). All NSF proposals are evaluated through use of the two merit review criteria. In some instances, however, NSF will employ additional criteria as required to highlight the specific objectives of certain programs and activities.

On July 8, 2002, the NSF Director issued Important Notice 127, Implementation of new Grant Proposal Guide Requirements Related to the Broader Impacts Criterion. This Important Notice reinforces the importance of addressing both criteria in the preparation and review of all proposals submitted to NSF. NSF continues to strengthen its internal processes to ensure that both of the merit review criteria are addressed when making funding decisions.

In an effort to increase compliance with these requirements, the January 2002 issuance of the GPG incorporated revised proposal preparation guidelines relating to the development of the Project Summary and Project Description. Chapter II of the GPG specifies that Principal Investigators (PIs) must address both merit review criteria in separate statements within the one-page Project Summary. This chapter also reiterates that broader impacts resulting from the proposed project must be addressed in the Project Description and described as an integral part of the narrative.

Effective October 1, 2002, NSF will return without review proposals that do not separately address both merit review criteria within the Project Summary. It is believed that these changes to NSF proposal preparation and processing guidelines will more clearly articulate the importance of broader impacts to NSF-funded projects.

The two National Science Board approved merit review criteria are listed below (see the Grant Proposal Guide Chapter III.A for further information). The criteria include considerations that help define them. These considerations are suggestions and not all will apply to any given proposal. While proposers must address both merit review criteria, reviewers will be asked to address only those considerations that are relevant to the proposal being considered and for which he/she is qualified to make judgments.

**What is the intellectual merit of the proposed activity?**
How important is the proposed activity to advancing knowledge and understanding within its own field or across different fields? How well qualified is the proposer (individual or team) to conduct the project? (If appropriate, the reviewer will comment on the quality of the prior work.) To what extent does the proposed activity suggest and explore creative and original concepts? How well conceived and organized is the proposed activity? Is there sufficient access to resources?

**What are the broader impacts of the proposed activity?**
How well does the activity advance discovery and understanding while promoting teaching, training, and learning? How well does the proposed activity broaden the participation of underrepresented groups (e.g., gender, ethnicity, disability, geographic, etc.)? To what extent will it enhance the infrastructure for research and education, such as facilities, instrumentation, networks, and partnerships? Will the results be disseminated broadly to enhance scientific and technological understanding? What may be the benefits of the proposed activity to society?

NSF staff will give careful consideration to the following in making funding decisions:

*Integration of Research and Education*
One of the principal strategies in support of NSF's goals 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 diversity of learning perspectives.

*Integrating Diversity into NSF Programs, Projects, and Activities*
Broadening opportunities and enabling the participation of all citizens -- women and men, underrepresented minorities, and persons with disabilities -- 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.

**Additional Review Criteria**:
- Possession of the scientific expertise and resources needed for tool development.
- Possession of the scientific expertise and resources needed for the creation and analysis of databases on organizations and individuals.
- Cohesion of technology, tools and data within each "testbed".
- Documented outreach and dissemination plan.
- Evidence of applicability to a broad range of sciences.
- Quality of coordination plan.
- Demonstration of scalability to, for example, additional organizations or other large-scale databases.
- Evidence of long-term sustainability and impact.

## B. Review Protocol and Associated Customer Service Standard

All proposals are carefully reviewed by at least three other persons outside NSF who are experts in the particular field represented by the proposal. Proposals submitted in response to this announcement wil be reviewed by an initial panel. The panelists will be asked to recommend either (1) further consideration by NSF or (2) no further consideration by NSF. Those teams whose proposals are recommended for further consideration will be asked to make a presentation to NSF program officers at NSF. After this "reverse site visit" to NSF, award recommendations will be made by NSF program officers .

Reviewers will be asked to formulate a recommendation to either support or decline each proposal. The Program Officer assigned to manage the proposal's review will consider the advice of reviewers and will formulate a 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 Director. In addition, the proposer will receive an explanation of the decision to award or decline funding.

NSF is striving to be able to tell proposers whether their proposals have been declined or recommended for funding within six months. The time interval begins on the closing date of an announcement/solicitation, or the date of proposal receipt, whichever is later. The interval ends when the Division Director accepts the Program Officer's recommendation.

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 Division 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.A. for additional information on the review process.)

## B. Award Conditions

An NSF award consists of: (1) the award letter, 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 letter; (4) the applicable award conditions, such as Grant General Conditions (NSF-GC-1); * or Federal Demonstration Partnership (FDP) Terms and Conditions * and (5) any announcement or other NSF issuance that may be incorporated by reference in the award letter. Cooperative agreement awards are administered in accordance with NSF Cooperative Agreement Financial and Administrative Terms and Conditions (CA-FATC). Electronic mail notification is the preferred way to transmit NSF awards to organizations that have electronic mail capabilities and have requested such notification from the Division of Grants and Agreements.

Consistent with the requirements of OMB Circular A-16, *Coordination of Geographic Information and Related Spatial Data Activities*, and the Federal Geographic Data Committee, all NSF awards that result in relevant geospatial data must be submitted to Geospatial One-Stop in accordance with the guidelines provided at: www.geodata.gov.

More comprehensive information on NSF Award Conditions is contained in the NSF *Grant Policy Manual* (GPM) Chapter II, available electronically on the NSF Website at http://www.nsf.gov/publications/pub_summ.jsp?ods_key=gpm. The GPM is also for sale through the Superintendent of Documents, Government Printing Office (GPO), Washington, DC 20402. The telephone number at GPO for subscription information is (202) 512-1800. The GPM may be ordered through the GPO Website at http://www.gpo.gov/.

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

## C. Reporting Requirements

For all multi-year grants (including both standard and continuing grants), the PI must submit an annual project report to the cognizant Program Officer at least 90 days before the end of the current budget period.

Within 90 days after the expiration of an award, the PI also is required to submit a final project report. Failure to provide final technical reports delays NSF review and processing of pending proposals for the PI and all Co-PIs. 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 FastLane, for preparation and submission of annual and final project reports. This system permits electronic submission and updating of project reports, including information on project participants (individual and organizational), activities and findings, publications, and other specific products and contributions. PIs will not be required to re-enter information previously provided, either with a proposal or in earlier updates using the electronic system.

## VIII. CONTACTS FOR ADDITIONAL INFORMATION

General inquiries regarding this program should be made to:

- Miriam Heller, Program Director, Office of the Director, Office of Cyberinfrastructure, 1145 S, telephone: (703) 292-7025, fax: (703) 292-9060, email: mheller@nsf.gov
- Lawrence E. Brandt, Program Manager, Directorate for Computer & Information Science & Engineering, Division of Information and Intelligent Systems, 1125 S, telephone: (703) 292-8930, fax: (703) 292-9073, email: lbrandt@nsf.gov
- Joan Maling, Program Director, Directorate for Social, Behavioral & Economic Sciences, Division of Behavioral and Cognitive Sciences, 995 N, telephone: (703) 292-8046, fax: (703) 292-9068, email: jmaling@nsf.gov
- Jacqueline Meszaros, Program Director, Directorate for Social, Behavioral & Economic Sciences, Division of Social and Economic Sciences, 995 N, telephone: (703) 292-7261, fax: (703) 292-9068, email: jmeszaro@nsf.gov
- Kevin L. Thompson, Program Director, Office of the Director, Office of Cyberinfrastructure, 1145 S, telephone: (703) 292-8962, fax: (703) 292-9060, email: kthompso@nsf.gov

For questions related to the use of FastLane, contact:

- Robbie W. Brown, Program Assistant, Directorate for Social, Behavioral & Economic Sciences, Division of Social and Economic Sciences, 995 N, telephone: (703) 292-7264, fax: (703) 292-9068, email: rbrown@nsf.gov

## IX. OTHER PROGRAMS OF INTEREST

The NSF *Guide to Programs* is a compilation of funding for research and education in science, mathematics, and engineering. The NSF *Guide to Programs* is available electronically at http://www.nsf.gov/cgi-bin/getpub?gp. General descriptions of NSF programs, research areas, and eligibility information for proposal submission are provided in each chapter.

Many NSF programs offer announcements or solicitations concerning specific proposal requirements. To obtain additional information about these requirements, contact the appropriate NSF program offices. Any changes in NSF's fiscal year programs occurring after press time for the *Guide to Programs* will be announced in the NSF E-Bulletin, which is updated daily on the NSF Website at http://www.nsf.gov/home/ebulletin, and in individual program announcements/solicitations. Subscribers can also sign up for NSF's MyNSF News Service (http://www.nsf.gov/mynsf/) to be notified of new funding opportunities that become available.

## ABOUT THE NATIONAL SCIENCE FOUNDATION

The National Science Foundation (NSF) funds research and education in most fields of science and engineering. Awardees are wholly responsible for conducting their project activities and preparing the results for publication. Thus, the Foundation does not assume responsibility for such findings or their interpretation.

NSF welcomes proposals from all qualified scientists, engineers and educators. The Foundation strongly encourages women, minorities and persons with disabilities to compete fully in its programs. In accordance with Federal statutes, regulations and NSF policies, no person on grounds of race, color, age, sex, national origin or disability shall be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving financial assistance from NSF, although some programs may have special requirements that limit eligibility.

*Facilitation Awards for Scientists and Engineers with Disabilities* (FASED) provide funding for special assistance or equipment to enable persons with disabilities (investigators and other staff, including student research assistants) to work on NSF-supported projects. See the GPG Chapter II, Section D.2 for instructions regarding preparation of these types of proposals.

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** (703) 292-5111
  (NSF Information Center)**:**

- **TDD (for the hearing-impaired):** (703) 292-5090

- **To Order Publications or Forms:**

    Send an e-mail to: pubs@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; 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 applicant 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 needing information as part of the review process or in order to coordinate programs; 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," 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records," 63 Federal Register 268 (January 5, 1998). 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 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 this burden estimate and any other aspect of this collection of information, including suggestions for reducing this burden, to: Suzanne Plimpton, Reports Clearance Officer, Division of Administrative Services, National Science

Foundation, Arlington, VA 22230.

OMB control number: 3145-0058.

Policies and Important Links | Privacy | FOIA | Help | Contact NSF | Contact Web Master | SiteMap

The National Science Foundation, 4201 Wilson Boulevard, Arlington, Virginia 22230, USA
Tel: (703) 292-5111, FIRS: (800) 877-8339 | TDD: (800) 281-8749

Last Updated:
06/09/05
Text Only