



National Science Foundation
4201 Wilson Boulevard
Arlington, Virginia 22230

Sustainable Digital Data Preservation and Access Network

Frequently Asked Questions

December 18, 2007

- Do the projects have to build a production model that is sustaining? Or just show handoff elsewhere for sustainability by doing appropriate research?
- The solicitation seeks to cover a broad range of activities. Is an alliance of existing organizations an acceptable path to shaping a DataNet Partner?
- Will the PowerPoint file from the DataNet Prospective Principal Investigator Informational Meeting held on November 6, 2007 be provided online?
- Please clarify the balance of domain science expertise and preservation and access/infrastructure in the ideal DataNet Partner. What do you see as a budget split between the efforts for infrastructure and CS research?
- Will cost share from the lead institution cause the proposal to be more likely supported?
- In listing “computer, computational, and information sciences” separately from “cyberinfrastructure,” does NSF mean to imply that these are separate areas of expertise requiring attention or can the proposer suggest an integrated computer science/information science/CI capability?
- What is the role of assessment and evaluation in DataNet?
- How broad does a repository need to be? ICPSR? Biology? Broader?
- How will all DataNet Partners self-assemble into a program-wide governance model? How should the proposal address this post-award task?
- What is data? The solicitation talks about scientific data as the “output of research.” Dan Atkins spoke of “digital objects of all sorts.” Are the following data: web sites, digital

books, journals, and photographs?

- What is the role of proprietary software and data/document formats?
- What objectives does DataNet have in mind with regard to educational outreach? What levels? Higher education? K-12? What forms of curriculum integration/ development? Faculty/student participation?
- Must all partners be named as part of the pre-proposal?
- Is research on data semantics within scope?
- Is there a limit on the number of proposals that an institution or individual may submit
- Is it necessary to name each proposed member of an advisory board at the pre-proposal stage? If so, is it necessary to provide bios and conflict of interest forms for each member?
- What evidence of participation for key personnel and partner organizations is wanted? For example, letters of support from the person or a senior administrator of their organization?
- Are any supplementary materials allowed beyond the references and information about senior and key personnel? For example, letters of support or other appendices?
- How should we handle the situation where it is not yet clear which individual from a partner organization would work with the project? Can we name an organization rather than a specific person? If we mention a person within a partner organization that is not a traditional research organization, how do we decide if they would be considered key personnel requiring a bio and conflict of interest form?
- Can additional partners - especially unfunded science partners still be added beyond the pre-proposal stage? More generally, how much flexibility is there to adjust the organization structure before the final submission?
- Similarly, if a funded commercial partner is identified during the project (for example, a third party storage provider) are there constraints? Must all recipients of project funding be named in the pre-proposal?

- Can changes be accepted to pre-proposals after they have been submitted to FastLane? For example, can we append additional bios or conflict of interest forms after initial submission but before the deadline?
- May we include international partners in our proposal? If so, can any part of their work be funded by NSF, especially if they work within our borders?
- May FFRDCs participate if they are subawardees?
- Can we include in the budget the cost of purchasing computers/servers, storage system and media, system operations/maintenance?

1. Do the projects have to build a production model that is sustaining? Or just show handoff elsewhere for sustainability by doing appropriate research?

The DataNet program aims to build sustainable infrastructure. Proposals that address only the research challenges without addressing the challenge of production quality infrastructure will be deemed not responsive to the solicitation and will be returned without review.

2. The solicitation seeks to cover a broad range of activities. Is an alliance of existing organizations an acceptable path to shaping a DataNet Partner?

DataNet seeks to create a new type of organization that we do not believe exists today. An existing alliance of organizations would have to clearly demonstrate in the proposal that they can meet all of the requirements of a DataNet Partner. Organizations seeking funding to continue their current business as usual, without significant re-thinking, should not seek DataNet funding.

3. Will the PowerPoint file from the DataNet Prospective Principal Investigator Informational Meeting held on November 6, 2007 be provided online?

Yes, the slides are available on the OCI website at (http://nsf.gov/news/news_summ.jsp?cntn_id=110392&org=OCI&from=news). They are also available directly by sending an e-mail request to DataNet@nsf.gov.

4. Please clarify the balance of domain science expertise and preservation and access/ infrastructure in the ideal DataNet Partner. What do you see as a budget split

between the efforts for infrastructure and CS research?

There is no one-size-fits-all answer to this question. We believe in user/user-centered design. We also believe that librarians, archivists, and computer/computational/information scientists are unlikely to build excellent infrastructure for science and/or engineering without deep engagement with the intended users. In that sense, domain scientists should be full partners in the process.

5. Will cost share from the lead institution cause the proposal to be more likely supported?

Cost sharing is not a requirement of the solicitation and it is not an evaluation criterion.

6. In listing “computer, computational, and information sciences” separately from “cyberinfrastructure,” does NSF mean to imply that these are separate areas of expertise requiring attention or can the proposer suggest an integrated computer science/information science/CI capability?

Integrated capability is acceptable. The point is that computer/computational/information science research is required to meet the goals of DataNet where we do not have existing solutions that are effective. That is distinct from the tasks involved in creating, managing, and sustaining cyberinfrastructure.

7. What is the role of assessment and evaluation in DataNet?

The project Description in the full proposal includes the following under Activities: “Provide plans for developing and implementing a vigorous and comprehensive assessment and evaluation program.” Assessment and evaluation is about measuring the degree to which a project has met or is meeting the intellectual and broader impact goals described in the proposal. It provides opportunities for course corrections that keep a project on target. More importantly, effective assessment and evaluation can make the case for the value proposition offered by a DataNet Partner and thus contribute to the goal of sustainability.

8. How broad does a repository need to be? ICPSR? Biology? Broader?

The initial size or range of disciplines/communities covered by a DataNet repository is less an issue than the degree to which the potential Partner addresses the challenges of being open, extensible and evolvable. A key concern is that infrastructure designed with one discipline in mind may not scale or adapt for

another, unless multiple disciplines were considered during the design phase and evolution was a recognized goal.

A DataNet Partner might start with a relatively small collection and user base that allowed it to address critical issues, with a clear plan and process for scaling and evolving to something much broader. However, breadth alone is not sufficient; a potential Partner might start with a very large collection and a large user base and still fail the tests of vision and planned strategy for evolution.

Vision and planning are both critical.

9. How will all DataNet Partners self-assemble into a program-wide governance model? How should the proposal address this post-award task?

DataNet Partners will be expected to collaborate with one another to develop a broad governance model. For example, one way to address this expectation is to budget for the collaboration, both in terms of travel and time. Another way is to build this social development process into the management plan. Applicants are encouraged to think creatively about how to meet this goal.

10. What is data? The solicitation talks about scientific data as the “output of research.” Dan Atkins spoke of “digital objects of all sorts.” Are the following data: web sites, digital books, journals, and photographs?

The DataNet solicitation defines data as “any information that can be stored in digital form and accessed electronically, including, but not limited to, numeric data, text, publications, sensor streams, video, audio, algorithms, software, models and simulations, images, etc.” Thus all of the objects mentioned in the question qualify as data provided they are in digital form and can be accessed electronically.

The DataNet applies the additional constraint that data must pass the test of relevance to the science and engineering research and education missions of NSF. Data that do not meet this constraint are out of scope for the DataNet program, though they may be of tremendous social and/or cultural value.

11. What is the role of proprietary software and data/document formats?

There exists a wide variety of commercial/proprietary software and data/document formats that are pertinent to DataNet. NSF’s expectation is that existing solutions will be used where they meet the scientific and economic needs of the project. New software, formats, etc. that are developed for DataNet with

government funding are required to be open source.

12. What objectives does DataNet have in mind with regard to educational outreach? What levels? Higher education? K-12? What forms of curriculum integration/development? Faculty/student participation?

Among the responsibilities of DataNet Partners, one is to provide systems, tools and resources that enhance the use of digital data for education and training purposes at all levels as well as capabilities for integration of research and education at all levels.

All proposal submitted to the NSF must be evaluated for their Broader Impact, which counts equally with Intellectual Merit. The Broader Impact criterion applied to DataNet proposals follows:

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? In addition, full proposal review criteria include the need to address developing new tools and capabilities for learning that integrate research and education at all levels.

13. Must all partners be named as part of the pre-proposal?

All partners who will receive funding from the project must be named in the pre-proposal.

14. Is research on data semantics within scope?

Research on data semantics is within scope to the extent that it is essential to fulfilling the requirements of a successful DataNet Partner.

15. Is there a limit on the number of proposals that an institution or individual may submit:

As stated in the DataNet solicitation, no limit is specified on the number of proposals per organization and no limit is specified on the number of proposals per PI.

16. **Is it necessary to name each proposed member of an advisory board at the pre-proposal stage? If so, is it necessary to provide bios and conflict of interest forms for each member?**

No, it is not necessary. See response to question 13.

17. **What evidence of participation for key personnel and partner organizations is wanted? For example, letters of support from the person or a senior administrator of their organization?**

Key personnel should provide curriculum vitae and conflict-of-interest information. Letters of collaboration from partner organizations are appropriate (but not required) for the full proposal. Letters of support are not part of the NSF proposal review process.

18. **Are any supplementary materials allowed beyond the references and information about senior and key personnel? For example, letters of support or other appendices?**

Letters of support are not part of the NSF proposal review process. See answer to question 17.

19. **How should we handle the situation where it is not yet clear which individual from a partner organization would work with the project? Can we name an organization rather than a specific person? If we mention a person within a partner organization that is not a traditional research organization, how do we decide if they would be considered key personnel requiring a bio and conflict of interest form?**

You can name the organization. Key people have key roles in management and leadership. You are limited by page space, so list the people who contribute to the vision and intellectual effort of the proposal. Also, see the response to question 13.

20. **Can additional partners - especially unfunded science partners still be added beyond the pre-proposal stage? More generally, how much flexibility is there to adjust the organization structure before the final submission?**

Unfunded partners can be added to the full proposal even if they are not listed in the preliminary proposal. Also, see the response to question 13.

21. **Similarly, if a funded commercial partner is identified during the project (for**

example, a third party storage provider) are there constraints? Must all recipients of project funding be named in the pre-proposal?

Known funded partners should be added if at all possible, since that is part of how the proposal gets reviewed. Also, see the response to question 13.

- 22. Can changes be accepted to pre-proposals after they have been submitted to FastLane? For example, can we append additional bios or conflict of interest forms after initial submission but before the deadline?**

Prior to the deadline announced in the solicitation, submitted proposals may be withdrawn and resubmitted, but the final version must be submitted by the deadline. See the NSF Grant Proposal Guide, Chapter I, Section F (http://www.nsf.gov/pubs/policydocs/pappguide/nsf08_1/gpg_1.jsp#IF) for further information regarding deadlines.

- 23. May we include international partners in our proposal? If so, can any part of their work be funded by NSF, especially if they work within our borders?**

As noted in the DataNet Solicitation, “NSF strongly encourages active international collaboration by the DataNet Partners. DataNet funds may be used to support U.S. investigators and students to work in international settings and foreign investigators and students to work in the U.S. However, foreign collaborators are expected to secure support for their activities at their home or other non-U.S. institutions from their own national sources.”

- 24. May FFRDCs participate if they are subawardees?**

Yes, subawards to FFRDCs are permitted.

- 25. Can we include in the budget the cost of purchasing computers/servers, storage system and media, system operations/maintenance?**

Yes. See the DataNet solicitation instructions for Appendix A3 for further information. Also see the NSF Grant Proposal Guide, Chapter II, Section C.2.g.iii (http://www.nsf.gov/pubs/policydocs/pappguide/nsf08_1/gpg_2.jsp#IIC2giii), which states:

“Equipment is defined as an item of property that has an acquisition cost of \$5,000 or more (unless the organization has established lower levels) and an expected service life of more than one year. It is important to note that the

acquisition cost of equipment includes modifications, attachments, and accessories necessary to make the property usable for the purpose for which it was purchased. Items of needed equipment must be adequately justified, listed individually by description and estimated cost.

Allowable items ordinarily will be limited to research equipment and apparatus not already available for the conduct of the work. General-purpose equipment, such as a personal computer and office furnishings, are not eligible for support unless primarily or exclusively used in the actual conduct of scientific research.

Additional information on the charging of equipment to an NSF award is available in the NSF Award and Administration Guide, [AAG Chapter V.B.2.](#)

As stated in the DataNet solicitation, “Funds for facility construction or renovation may NOT be requested.”