



# Sustainable Digital Data Preservation and Access Network (DataNet)

**Lucy Nowell, Ph.D.**

**Program Director**

**Data, Data Analysis, & Visualization**

**Office of Cyberinfrastructure**

**Sylvia Spengler, Ph.D**

**Program Director**

**Information & Intelligent Systems**

**CISE**

**Webcast Dial-in: 1-877-624-8532. Passcode 281913**

**November 6, 2007**



# Today's Topics



- **DataNet solicitation - vision and goals**
- **DataNet Partner characteristics and activities**
- **Proposal requirements**
  - Preliminary proposal
  - Full proposal
- **Evaluation criteria**
- **Proposal review process**



# Vision



- **“Science and engineering digital data are routinely deposited in a well-documented form, are regularly and easily consulted and analyzed by specialists and non-specialists alike, are openly accessible while suitably protected, and are reliably preserved.”**
  - **NSF Cyberinfrastructure Vision for 21st Century Discovery**
- **Catalyze development of a system of science and engineering data collections that is open, extensible and evolvable.**



# Grand Challenge



- **Develop new methods, management structures and technologies to manage the diversity, size, and complexity of current and future data sets and data streams.**



# DataNet Goals

- A robust and resilient national and global digital data framework for **preservation and access** to the resources and products of the digital age
  - Provide reliable digital preservation, access, integration and analysis capabilities for science and/or engineering over a decades-long timeline: **sustainability**
  - Continuously anticipate and **adapt to changes** in technologies & user needs and expectations
  - Engage at the **frontiers** of science & engineering research & education, with **research** & development to drive the leading edge forward
  - Serve as component elements of an interoperable data preservation and access network, spanning national and international boundaries: **shared governance and standards**
- Creation of new types of organizations that fully integrate all of these capabilities



# Digital Data

- **Any information that can be stored in digital form and accessed electronically: numeric data, text, publications, sensor streams, video, audio, algorithms, software, models and simulations, images, etc.**
- **Caveats**
  - **Focus on data central to the scientific and engineering research and education mission of NSF.**
  - **Conversion to digital format is not supported by this program.**
  - **No support for narrowly defined, discipline-specific repositories.**



# DataNet Partner Requirements



- **Each exemplar DataNet Partner will**
  - **Combine expertise** in library and archival sciences; computer, computational and information sciences; cyberinfrastructure; and domain sciences and engineering
  - Develop models for **economic and technological sustainability over multiple decades**
  - Engage at the **frontiers of science and engineering research and education**
  - Work cooperatively and in coordination to to create a functional data network with **revolutionary new capabilities for information access, use, and integration** without regard to conventional barriers such as data type and format, discipline or subject area, and time and place/institution.



# DataNet Partner Characteristics



- **Risk averse & risk capable – provide stability for long-term preservation and agility both to embrace constant technological change & to engage in evolving research challenges**
- **Sustainability – economic models & technology strategies to meet decades-long mission for preservation & access**
- **Engage at the frontiers of research & education**
  - **Close interaction with relevant domain science communities to understand needs, practices & expectations**
  - **Facilitate research in digital concepts & technologies that support long-term preservation & access for complex digital objects & assets.**
- **Heterogeneous in content & collectively interoperating – support repurposing of digital data in innovative ways & novel combinations not envisioned by the data authors**
  - **Serve a broad disciplinary & subject matter range**
  - **Manage a diverse array of data types & formats**
  - **Support collections at research, resource, & reference levels**





# DataNet Partner Responsibilities



## 1. Vision and Rationale

- **Identify and meet specific scientific needs & create new capabilities for discovery, innovation, and learning**
- **Develop and implement sustainable economic and technology models**



# DataNet Partner Responsibilities



## 2. Activities

- **Provide for full data management life cycle**
  - Data deposition/acquisition/ingest
  - Data curation & metadata management
  - Data protection, including privacy
  - Data discovery, access, use, & dissemination
  - Data interoperability, standard, & integration
  - Data evaluation, analysis, & visualization
- **Engage in research central to DataNet responsibilities**
- **Education & training**
- **Community & user input assessment**
- **International engagement – collaborate & coordinate closely with preservation & access organizations to catalyze formation of a global data network**
  - Foreign collaborators are expected to secure support from their own national sources.



# DataNet Partner Responsibilities



## 3. Organizational Structure

- Comprehensive expertise & infrastructure capacity/capabilities – integrate expertise in library & archival sciences; computer, computational, & information sciences; cyberinfrastructure; & domain science/engineering
- Diverse, multi-sector participation – diverse academic organizations; federal, state & local governmental organizations; commercial entities; & international organizations
  - Develop policies & practices that encourage participation of women & underrepresented groups at all levels & in all Partner activities
- Data Network – collaborate & coordinate closely with other DataNet Partners & other preservation/access organizations to develop & disseminate best practices & principles; required development of shared governance or coordination framework



# DataNet Partner Leadership & Management



- **Director (not necessarily PI, may be TBD)**
  - Strong central leadership to insure unified vision & strategy
  - Broad, integrated vision; familiarity with diverse areas; ability to lead interdisciplinary teams
- **Effective management plan with clear lines of responsibility & authority**
- **Appropriate oversight boards & user groups**



# DataNet Awards



- **Award type: Cooperative Agreement**
- **2-3 awards in each of 2 review cycles**
- **Up to \$4,000,000 per year for up to 5 years (total \$20,000,000) for each of up to 5 awards**
- **Possible renewal for an additional 5 years, with declining funding each year, as Partners become self-sustaining**



# DataNet Eligibility



- **Eligible lead institutions:**
  - Academic institutions in the U.S.
  - Non-profit, non-academic organizations in the U.S. - independent museums, observatories, research labs, professional societies, and similar organizations with associated educational or research activities
- **Sub-awards may be made to all organizations eligible under the provisions of the NSF Grant Proposal Guide (GPG)**
- **Collaborating organizations must submit a **single proposal from the lead institution** and may NOT be submitted as linked, collaborative proposals.**



# Proposal Timeline



- **Preliminary Proposals (required) due**
  - **January 7, 2008**
  - **October 6, 2008**
  
- **Invited Full Proposals due**
  - **March 21, 2008**
  - **February 16, 2009**



# Preliminary Proposal



- **FastLane Cover Sheet**
- **Project Title must begin “DataNet Preliminary Proposal:”**
- **Project Summary (1 page) in 3 parts**
  - 1. Title, lead organization, PI , Co-PIs, & senior personnel**
  - 2. Intellectual Merit:** 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 prior work.) To what extent does the proposed activity suggest and explore creative, original, or potentially transformative concepts? How well conceived and organized is the proposed activity? Is there sufficient access to resources?
  - 3. Broader Impact:** 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?





# Preliminary Proposal



- **Project Description (7 pages)**
  - **Vision & rationale** – nature of needs & opportunities to be met & expected impact; overview strategy for achieving long term economic & technological sustainability
  - **Activities** to manage the full data life cycle, enabling discovery, innovation & learning with participation across diverse sectors
  - **Overview of structure** – sectors involved, nature of user base, diversity of participants & users, range of expertise & infrastructure to be included



# Preliminary Proposal



- **References cited**
- **Biographical sketches**
- **Current & pending support**
- **Budget not required at this stage - enter \$2 in Requested Amount box**
- **Special Information & Supplementary Documentation: Key leadership personnel (3 pages), with brief description of unique contribution of each**
- **Single Copy Documents (confidential):**
  - **Integrated conflict-of-interest list**
  - **Optional list of potential reviewers**



# Full Proposals (By Invitation Only)



- **Eligible proposals must originate from principal investigators whose proposals are successful in the preceding preliminary proposal competition described above and **must** be based on those preliminary proposals.**
- **Full proposals submitted by PIs whose preliminary proposals received a review recommendation of “not invited” will be returned without review.**



# Full Proposals



- **Cover Sheet - Submit via FastLane or Grants.gov**
- **Title must begin “DataNet Full Proposal:”**
- **Project Summary (1 page) in 3 parts**
  - 1. Title, lead organization, PI , Co-Pis, & senior personnel**
  - 2. Intellectual Merit: 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 prior work.) To what extent does the proposed activity suggest and explore creative, original, or potentially transformative concepts? How well conceived and organized is the proposed activity? Is there sufficient access to resources?**
  - 3. Broader Impact: 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?**



# Full Proposals



- **Project Description (15 pages)**
  - Results from prior research
  - Vision & rationale: Describe why the DataNet Partner is needed, new opportunities created, impact, rationale for choice of science/engineering fields served, anticipated user base, strategy for achieving long-term economic & technological sustainability
  - Activities: Plans for supporting the full data preservation and access life cycle, engagement at frontiers of science, education & training plans & capabilities for integrating research & education, plans for assessment & evaluation
  - Structure: How components will work together to achieve the vision. Overview management plan. Outline cyberinfrastructure components & how integrated into effective framework. Types of expertise provided w/contributions of each. Sectors involved, w/contributions & benefits for each. Plans for increasing participation of women & underrepresented groups in all Partner activities. Plans for promoting formation of a functional data network.



# Full Proposals



- **Project Description (15 pages) (cont.)**
  - References cited
  - Biographical sketches
  - Current & pending support
  - Budget
    - No funds for facility construction or renovation
    - Must include cost for required participation by 4 project personnel in annual meeting of DataNet leadership at NSF.
  - **Special Information & Supplementary Documentation**
    - Appendix A1, Sustainability Plans (4 pages)
    - Appendix A2, Management Plan (3 pages)
    - Appendix A3, Cyberinfrastructure Capabilities (4 pages)
    - Appendix A4, Key Personnel & Unique Contributions
    - Appendix A5, Role of Participating Organizations & Sectors (3 pages)
    - Appendix A6, Results from Prior Research (4 pages)



# Full Proposals



- **Single Copy Documents**
  - **Integrated Conflict-of-Interest List for PI, Co-PIs, & any named personnel whose salary is requested in project budget**
  - **Correspondence not to be sent to reviewers, such as optional list of potential reviewers**



# NSF Merit Review Criteria



- **Intellectual Merit:**
  - 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? 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?
- **Broader Impact:**
  - How well does the activity advance discovery and understanding while promoting teaching, training, and learning? How well does the proposed activity broaden participation of underrepresented groups? 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?





# Additional Review Criteria: Preliminary Proposals



- **Vision & Rationale:** How effective is the preliminary proposal in describing a compelling vision to meet important needs & opportunities & a credible strategy to achieve long term sustainability?
- **Activities:** How effective will the activities outlined in the preliminary proposal be in providing for management of the full data life cycle & enabling discovery, innovation, & learning with participation across diverse sectors?
- **Organizational Structure:** Does the structure outlined in the preliminary proposal provide for participation by an appropriate range of sectors, serve and engage an appropriate set of users & participants, & provide for necessary expertise & infrastructure?



# Additional Review Criteria: Invited Full Proposals



- **Vision & Rationale:** How effective will the plan described in the proposal be in
  - Meeting well defined & critical needs;
  - Creating new opportunities & capabilities for discovery;
  - Improving the way science & engineering research & education are conducted; and
  - Achieving long-term economic & technological sustainability?
- **Activities:** How effective will the plan described in the proposal be in
  - Providing for the full data management life cycle;
  - Engaging in research central to DataNet responsibilities & facilitating research as a resource & object;
  - Developing new tools & capabilities for learning that integrate research & education at all levels;
  - Providing for community input & participation in all phases & aspects of Partner activities; and
  - Ensuring vigorous & comprehensive evaluation & assessment of all aspects of the project?



# Additional Review Criteria: Invited Full Proposals



- **Organizational Structure: How effective will the plan described in the proposal be in**
  - Providing a comprehensive & appropriate range of expertise in library & archival sciences; computer, computational, & information sciences; cyberinfrastructure; & domain sciences?
  - Providing the required cyberinfrastructure resources & capabilities;
  - Serving a diverse user base;
  - Ensuring active participation by a diverse range of individuals (including women & underrepresented groups), organizations, & sectors;
  - Serving as an effective partner in an interoperable network of digital preservation & access organizations;
  - Providing a management plan for effective leadership with clear lines of authority, responsibility, accountability, community & user responsiveness, & the ability to adapt to new opportunities?



# Review Process



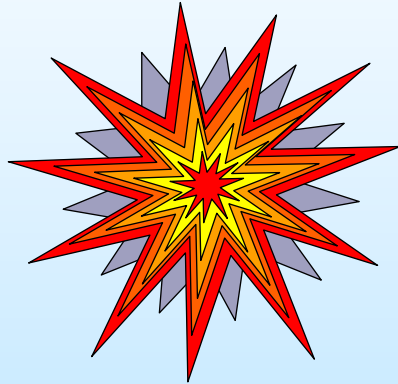
- **Review in 3 stages**
  - 1. Preliminary Proposals reviewed by panel of external reviewers. Successful preliminary proposals result in invitation to submit full proposal.**
  - 2. Invited Full Proposals reviewed by external mail-in and panel reviewers.**
  - 3. Site Visits to selected candidates result in panel recommendations regarding awards.**
- **Program Officers consider advice of reviewers and formulate recommendations to NSF.**



# Proposal Preparation Checklist



- Proposal submitted as single proposal with one organization as lead & other participating organizations as sub-awards; **NO linked collaborative proposals accepted.**
- Title begins “DataNet Preliminary Proposal:” or “DataNet Full Proposal:”
- Project Summary contains all 3 parts, including broader impacts of proposed work
- Project Description is 7 (preliminary) or 15 (full) pages or less, including figures & tables
- References Cited includes publications resulting from prior research funded by NSF (marked \*)
- Biographical Sketches (2 pages each) for PI, CO-PIs, & Senior Personnel listed in Project Summary
- Current and Pending Support Statements for PI, CO-PIs, & Senior Personnel
- For full proposal, required appendices A1, A2, A3, A4, A5 & A6 uploaded in Supplementary Documents
- Single, alphabetized listing of conflicts of interest uploaded into Single Copy Documents



Thank you!