National Science Foundation
4201 Wilson Boulevard
Arlington, Virginia 22230

Earth Cube Guidance for the Community

1. **EarthCube—Building the Cyberinfrastructure to Better Understand Our Complex and Changing Planet**: This document describes the GEO-OCI vision for EarthCube and the timeline of development activities.

2. **EarthCube Charrette Preparation**: This document describes the EarthCube Charrette process, which will include open community forums to identify requirements, stimulate discussion and foster partnerships and collaboration.

3. **EarthCube Charrette Interest in Attending**: We ask that individuals who are interested in attending the Charrette notify us of their interest and provide some information on their expertise and ideas on EarthCube.

EarthCube -- Building the Cyberinfrastructure to Better Understand Our Complex and Changing Planet

The intent of this document is to provide a context and vision for EarthCube. This is an open document to inform, stimulate discussion, and engage the community in building EarthCube.

Context

In 2009 the Advisory Committee for NSF's Geosciences Directorate (GEO) issued a report (*GEO Vision*) that identifies the challenges and opportunities facing the geosciences. The report identifies "*fostering a sustainable future through a better understanding of our complex and changing planet*" as the pre-eminent challenge and opportunity for the geosciences community. To accomplish this end, the report issued a call to action:

> "*Over the next decade, the geosciences community commits to developing a framework to understand and predict responses of the Earth as a system— from the space-atmosphere boundary to the core, including the influences of humans and ecosystems.*"

In 2011 NSF announced Cyberinfrastructure for the 21st Century (CIF21), an important new research thrust for FY 2012 led by the Office of Cyberinfrastructure (OCI). Within CIF21 significant emphasis is placed on computational and data-rich science and engineering. The vision of CIF21 is to provide the nation with a sustainable, community-based and open cyberinfrastructure for researchers and educators.

In response to the challenges and opportunities presented in GEO Vision, GEO and OCI have launched EarthCube. EarthCube will be the transformative vehicle to address the scientific drivers in the GEO Vision document: 1) Understanding and forecasting the behavior of a complex and evolving Earth system; 2) Reducing vulnerability and sustaining life; and 3) Growing the geosciences workforce of the future. EarthCube is inspired by the vision of a national cyberinfrastructure in CIF21. As an effort to build a unified cyberinfrastructure framework for the geosciences, EarthCube will be supported by CIF21 but will also help to define CIF21.

What is EarthCube

The goal of EarthCube is to create a knowledge management system and infrastructure that integrates all geosciences data in an open, transparent and inclusive manner. No integrated framework currently exists to meet the information management challenges implicit in the GEO Vision call to action. EarthCube is this goal, as well as a process that will require years of development, investment, and community engagement. The decade-long vision for EarthCube is the convergence towards an integrated system to access, analyze and share information that is used by the entire geosciences community. This convergence can only occur through community dialogue and collaboration. NSF will facilitate this convergence through currently planned and future activities.

Realizing EarthCube requires building a community derived and governed cyberinfrastructure that can be easily adopted by geosciences researchers and educators. It will require the introduction of new approaches and technologies and/or combining productive tools and solutions in different ways. EarthCube will promote integration, flexibility, inclusiveness, and easy adoption by connecting the several layers of data and information management, from the resource layer with access to data and information, to the data curation and management layer, and to the overarching interface layer that exposes data and information to knowledge creation through data-enabled science. Interworkability of data and information will be a characteristic of EarthCube, which implies interoperability and sharing, but with a human-centric focus that allows researchers and educators to easily work with multiply layers of technology and information. EarthCube will create new modes of learning and training that result in a more informed public and policy-makers while simultaneously broadening participation in the creation of a sustainable Earth system.

EarthCube will be supported by the substantial cyberinfrastructure investments, including databases, software services and community facilities, that have been created by the geosciences community over the past two decades. The community must build on this existing foundation of cyberinfrastructure and community knowledge to create an *integrated* set of services serving the entire geosciences community.

The success of EarthCube will depend in part on identification of commonalities of solutions and best practices that reside within our current infrastructure and strategic adoption of new technologies and approaches external to the community's infrastructure. All geosciences cyberinfrastructure will evolve over time to accommodate changing user needs and emerging technologies and services. Results from the community-driven EarthCube effort will provide an architectural framework to guide the integration of efforts as well as the evolution of existing cyberinfrastructure.

Assumptions

NSF makes several assumptions in the effort to build EarthCube. These are:

- The geosciences community is ready to take on the EarthCube challenge

- There is an existing infrastructure and knowledge base on which EarthCube will be built, but there is more to do in order to build an integrated framework

- In the next decade, technology will continue to evolve at a pace that will allow for the convergence and integration of systems expected under EarthCube

- The community is willing to come together in a process marked by distinct events and virtual and real dialog about all or parts of EarthCube. These opportunities for dialog and collaboration will be facilitated by NSF though presentations, social media websites, and community meetings. The first community meeting, the EarthCube Charrette, will be a collaborative session in which possible solutions for an EarthCube design are drafted by stakeholders in the community.

- The challenge will not be easy or rapidly addressed. However, it is expected that there will be a convergence to a common framework over time and that the EarthCube Charrette is the beginning of this process

- Initial self-organizing efforts will occur prior to the EarthCube Charrette. Groups and individuals will use an online forum to be provided to create partnerships and share their ideas prior to the Charrette

- A broad range of expertise will be required during the Charrette. Representation from the following groups is deemed important:

  o Geoscientists who have visionary knowledge of geosciences fields and research

  o Users who have a strong grasp of both the geosciences community's scientific and cyberinfrastructure needs

  o Cyberinfrastructure architects, builders and other technologists

  o Knowledge management/Information systems experts who can contribute to discussions related to turning user and data requirements into cyberinfrastructure functionality

o Individuals who have experience with the governance of multi-user infrastructure and have engaged the community in the creating, building, maintaining, and modifying facilities.

NSF expects, and will make efforts to ensure, that these categories of expertise will be well represented throughout the process described herein. Participation by early-career researchers, who may be most affected by EarthCube, is encouraged, as is the participation of international partners grappling with similar issues in their communities.

EarthCube Process

To create EarthCube, NSF is embarking on a process to help focus the best ideas and foster partnerships among a number of on-going geosciences activities and those activities in other fields that can contribute to advancing geosciences cyber-infrastructure. The process is a continuing dialog with the community marked by two focusing events: an EarthCube Charrette and a second EarthCube Community Event. These events will be intense periods of community-driven, interactive development of the partnerships, concepts, architecture and infrastructure that will form EarthCube. The EarthCube process begins with the following steps:

On-line Community Information: **(August to November, 2011)** to identify and make available community requirements; to develop partnerships that encompass the broad geosciences vision of EarthCube. To facilitate this process, resources will be available to view other plans, user requirements, etc.

EarthCube Charrette: (Early November, 2011) to present, refine, and integrate the best ideas, technologies, and approaches that are produced, with a focus on meeting the science needs of the community.

Post Charrette: (Mid-November to April, 2012) to codify community-influenced designs, governance structures, etc., of innovative approaches to EarthCube. EAGER grants may be supported for development and design of transformative concepts.

A Second EarthCube Community Event: **(Tentatively Early May, 2012)** to further the design development through a process of intense design and revision with the intention of funding prototype systems.

Prototype Development: (May to December 2013) to develop system capabilities, usefulness to a broad range of geosciences users and to create community governance founded on trust and common values.

Beyond the Prototype: (2014-2022): to continue support for the development of the best approach to EarthCube, guided by the milestones and deliverables established during the prototype development phase.

EarthCube -- Preparing for the EarthCube Charrette

The intent of this document is to provide NSF's expectations for the EarthCube Charrette and guidance to the community on preparing for the event.

The EarthCube process begins with a period of Community Organization (August-November) prior to the EarthCube Charrette (November 1-4, 2011). Following a webinar on August 22, 2011, NSF will facilitate the Community Organization by providing an online resource for gathering community inputs, including a forum for developing partnerships and a process for submitting EarthCube design approaches. This forum will be available in early September at the following address http://earthcube.ning.com. Also available will be a process for expressing interest in attending the Charrette. Space is limited at the Charrette venue, so groups are encouraged to self-organize and choose representatives. Virtual participation in the Charrette will not be limited.

NSF encourages communities to become involved early in the open Community Organization and to choose community representatives to participate in the Charrette. NSF expects a broad and diverse participation in the activities to build EarthCube, as noted in the EarthCube white paper: "EarthCube—Building the Cyberinfrastructure to Better Understand Our Complex and Changing Planet." Representation from the following groups of expertise is important:

- Geoscientists who have visionary knowledge of geosciences fields and research

- Users who have a strong grasp of both the geosciences community's scientific and cyberinfrastructure needs

- Cyberinfrastructure architects, builders and other technologists

- Experts in knowledge management/information systems who can contribute to discussions related to turning user and data requirements into cyberinfrastructure functionality

- Individuals who have experience with the governance of multi-user infrastructure and have engaged the community in creating, building, maintaining, and modifying facilities.

- Postdocs and Graduate Students who have training in the geosciences or in a technology field with an interest and ability to participate in discussions related to geosciences cyberinfrastructure.

NSF expects and will make every effort to have these categories of expertise well represented at the EarthCube Charrette. Participation of early-career researchers, who may be most affected by EarthCube, is encouraged as is participation by international partners grappling with similar issues in their communities.

Inputs Requested Prior to the EarthCube Charrette

EarthCube's process and outcomes will only be successful through open community dialog. Every effort will be made to broadly engage the scientific community at all points in the process.

Innovative solutions may come from individual researchers with insight into the geosciences or cutting edge technology solutions, as well as from institutions and organizations that support the geosciences, such as observational facilities, consortia representing scientific groups, and groups that manage cyberinfrastructure for various communities. Because partnerships and collaborations among these various groups are necessary for the broad geosciences-wide vision of EarthCube, NSF will provide online resources prior to and after the Charrette to facilitate dialog and partnership building. Using the online resources is not required to participate in the EarthCube process, and groups can self-organize as they choose. However, NSF reaffirms that EarthCube can only succeed through openness and sharing of ideas across the traditional boundaries of geosciences research.

The online forum will be a means for a diverse set of community groups and consortia, researchers, and educators to share ideas, introduce concepts, and find and develop collaborative efforts (see http://earthcube.ning.com). This on-line forum will allow submissions in three categories: User Requirements, Technology Solutions, and EarthCube Design Approaches. Submissions to the website will be open to all.

User Requirements

Understanding the user's scientific requirements and unmet cyberinfrastructure needs is a required step in building EarthCube and advancing the frontiers of knowledge in the geosciences. Some requirements might take the form of automation of mundane, but time consuming, tasks such as research data management. Other requirements may look to improve the scientific discovery process by making information more easily organized and accessible. User requirements might also envision the innovative science that could be possible in 10 years with an integrated cyberinfrastructure framework.

The on-line forum will include a questionnaire to be answered by the research community about their scientific focus areas or disciplines. These questions are presented in the appendix. The response from the community will be important input to test the scope and functionality of proposed approaches to the EarthCube system. Other user requirements, or use cases, already compiled by various geosciences

communities are welcome additions, and may also be uploaded to the forum. During the Charrette, this input will complement the knowledge of participants. Respondents to the User Requirements Questions may be invited to the Charrette to share their perspectives on the challenges and scientific questions in the geosciences.

User requirements are dynamic and will be needed to be continually assessed over time. The process to build EarthCube, as well as the system itself, will have to accommodate new data sources and user requirements.

Technology Solutions

Communities within geosciences, as well as other sciences, have created technology-based solutions to manage and use information pertinent to their discipline. These solutions could become important elements of EarthCube.

NSF encourages groups to use the online forum to share their innovative solutions to data management, data analytics, and any aspect of cyberinfrastructure that may support scientific processes important to the geosciences. These solutions may come from smaller communities within the geosciences that have self-organized. Some may be advanced by those in industry, academia, international groups and federal agencies. Technology solutions may be new, relatively untested ideas as well as frameworks and infrastructure solutions that have worked well for different constituencies.

Technology solutions should be uploaded to the on-line forum in short white-papers, no more than 5-6 pages (including appendices). These should describe

- The vision for what they wanted to solve/create
- How they defined and addressed user requirements
- How they partnered with other entities
- CI architecture design, development and integration
- Future plans, including further development needs and outreach to a broader geosciences community

EarthCube Design Approaches

Finally, the community is invited to share their ideas for the design and implementation of EarthCube. Groups or individuals may submit initial position papers to the online forum that discuss their ideas and solicit input from the community. These ideas do not have to be completely fleshed out either in scientific scope or in cyberinfrastructure solutions. Early submission is encouraged to allow time to strengthen approaches through community dialog and new collaborations. The EarthCube Charrette will also be an opportunity for the community to learn about EarthCube Design Approaches, find

additional partnerships, expand scientific scope or adopt new solutions to develop more complete approaches.

Shortly before the Charrette, groups that would like their designs considered and significantly discussed and modified at the Charrette will need to submit a written document: an Earth Cube Design Approach. This document should be no more than 10 pages (including attachments and appendices), should comprehensively describe the elements of the approach that are relatively mature as well as those that need addition input and consideration. The document should acknowledge how the approach may be incomplete and how it will be open to engage other groups. The document should be readable and understood by a broad audience of researchers. EarthCube Design Approaches need to address the following elements:

- Vision for EarthCube
  - A description of the envisioned scope of EarthCube and how will it transform geosciences research, including the functionality EarthCube can provide to the whole Geosciences community.
- Community-Based Governance model
  - The community structures necessary to acquire current and future user input/requirements, to respond to changing data and science needs, to adapt and adopt new technologies, to coordinate components and facilities, to foster partnerships and community participation.
- Conceptual CI Architecture
  - The architecture necessary to provide the services of EarthCube, to integrate advanced information technologies that facilitate access to distributed resources such as computational tools and services, instruments, data, and people.
- Design Process
  - User requirement-driven design methodology, identification of design team members, qualifications of development team, time-line for design demonstration and scale-up, design tools and practices that create robust, sustainable, well-documented and open source infrastructure
- Operations and Sustainability Model
  - Operational aspects of a community-wide enterprise that address such activities as centralized functions, coordination of services, user services, including training, and identification of what it will take to sustain a viable infrastructure over a long periods of time and who will carry out these functions.

These submissions will be considered by NSF, and submitters will be contacted and offered the opportunity to present at the Charrette. While incomplete EarthCube Design

Approaches are expected, NSF will look first to those design approaches that make an effort to address all five elements and seriously consider and explain how the approach will address all of the geosciences. The Charrette process will provide an opportunity for further partnerships to develop that incorporate innovative approaches, or combine similar visions and plans, and ultimately lead to a strengthened and more broadly based design.

Process for Attending the EarthCube Charrette

NSF recognizes that for a successful outcome to the EarthCube Charrette there is a practical limit to the number of attendees. The intent is to have diverse participation from the many types of users, innovators, and potential design groups. Therefore, limits may be placed on the number of participants from the same institution, organization, area of expertise, or community as defined by some other criteria that may be applied. It is important that the ideas from the various communities be represented at the Charrette. Thus, it is important that communities, working groups, university departments, non-governmental organizations, etc., appoint representative individuals to attend the Charrette.

For planning purposes, NSF requests that individuals notify NSF of their interest in attending the EarthCube Charrette. Instructions for this notification are contained within the document "EarthCube—Interest in Attending the EarthCube Charrette". This document will be available online. Individuals should respond as instructed in the document before **October 5, 2011**. First priority in the registration process will be given to those who respond before this date. The formal registration process, including criteria to balance participation, will be made available to the community early in October, 2011.

What to Expect at the EarthCube Charrette

The EarthCube Charrette is the first stage in an iterative process to build a community-based cyberinfrastructure. It is also a new process for the NSF to engage the geosciences community. Thus, open community dialog and responses to NSF's request for input will help to define the structure of the Charrette and the necessary discussions at the meeting. For example discussions might include broad science questions that define data requirements, new technologies, different governance models and design approaches.

The Charrette will be an opportunity for face-to-face interaction and feedback among groups that are suggesting designs for EarthCube and a broad, diverse set of the geosciences and cyberinfrastructure communities. Remote participation and interaction in the Charrette sessions will be provided via web-based resources.

The intention of the Charrette is to identify and focus the most useful and innovative ideas to meet the broad geosciences vision of EarthCube. It is important that groups come to the Charrette open and willing to change or adapt their ideas. It is possible that

groups and design ideas may merge, and/or that designs will become refined through real-time comments and breakout sessions. There will be opportunities to listen to other user requirements, innovative technological solutions, and proposed design approaches.

The process relies upon active involvement in these sessions by all participants. NSF and the facilitator will look for common solutions and encourage collaborations between groups. After the Charrette, groups would be free to consider submission of EAGER proposals to NSF. Attendance or participation in the EarthCube Charrette will not be a requirement for participating in any other steps of the EarthCube process.

User Requirements Questions

This survey is entirely voluntary. The feedback received here may be made available to the geosciences community to use in design development of EarthCube. You may decide whether you wish to include your name and affiliation with your response.

1. Briefly describe the scientific questions you are trying to address in your current work, and hope to address in 5 years, in 10 years.
2. What data and products do you need to answer these science questions?

    2a. Do you require real-time, digital, and/or physical data or data products? Does/will this data come from observing systems supported by the NSF, or other agencies?

    2b. Do you require data or tools from outside your immediate discipline? Name or describe the fields from which you would want data.

    2c. Are there sufficient tools (software, computational resources, visualization, etc. tools) to analyze your data?
3. Are you/ will you able to easily obtain and use the data and tools you need? If so, are there improvements that would facilitate your work? If not, what is lacking? Is there a place for products of your own work to be archived and accessed?
4. What tasks do you currently spend time on that you believe could be automated?
5. How important is cyberinfrastructure (CI) to the conduct of your research? What type CI do you currently use?
6. Do you have needs that you believe are common to numerous geosciences researchers? What are needs that are unique to your research?
7. Do you have other comments related to the cyberinfrastructure needed for your research?

EarthCube – Interest in Attending the EarthCube Charrette

NSF is interested in the number of attendees expected at the EarthCube Charrette and recognizes that for successful outcomes, there is a practical limit to the number of attendees. To facilitate planning this meeting, NSF requests that individuals interested in attending the EarthCube Charrette please submit answers to the following questions at this website http://www.surveymonkey.com/s/S8YCP8B. Individuals who respond before **October 5, 2011** will be given priority in the registration process. Other criteria may be needed to limit registration and will be made available in early October, 2011.

1. What is your name, affiliation, and academic or business contact information?

2. Please identify up to two groups of expertise, as listed in the document "EarthCube-Preparing for the EarthCube Charrette," that best describe you. Groups of expertise are:

o   Geoscientists who have visionary knowledge of geosciences fields and research

o   Users who have a strong grasp of both the geosciences community's scientific and cyberinfrastructure needs

o   Cyberinfrastructure architects, builders and other technologists

o   Experts in knowledge management/information systems who can contribute to discussions related to turning user and data requirements into cyberinfrastructure functionality

o   Individuals who have experience with the governance of multi-user infrastructure and have engaged the community in creating, building, maintaining, and modifying facilities.

o   Postdocs and Graduate Students who have training in the geosciences or in a technology field with an interest and ability to participate in discussions related to geosciences cyberinfrastructure.

Participation in the EarthCube online activities and forums is optional, but encouraged: http://earthcube.ning.com.