EarthCube Roadmap Document

DRAFT - 20 January 2016

Prepared by:

<u>EarthCube Leadership Council</u>

David Arctur
At-Large Representative
University of Texas, Austin

Marjorie Chan
Engagement Team Representative
University of Utah

Yolanda Gil
Technology and Architecture Committee Representative
USC's Information Sciences Institute

Basil Gomez
Science Committee Representative
University of Hawaiʻi, Mānoa

Farzad Kamalabadi
At-Large Representative
University of Illinois, Urbana

Danie Kinkade
At-Large Representative
Woods Hole Oceanographic Institution

Kerstin Lehnert
Chair
Columbia University

Lindsay Powers
Liaison Team Representative
The HDF Group

Mohan Ramamurthy
Council of Data Facilities Representative
Unidata, UCAR

Lynn Yarmey
At-Large Representative
National Snow & Ice Data Center

<u>EarthCube Project Office</u>

Lee Allison
PI, EarthCube Test Governance
Arizona Geological Survey

Cindy Dick
Project Office Manager
Arizona Geological Survey

Chris Lenhardt
Technical Lead
Renaissance Computing Institute, UNC-CH

EarthCube Roadmap Document

DRAFT - 20 January 2016

I. Introduction

EarthCube is a holistic approach to advancing cyberinfrastructure for the geosciences toward the ultimate goal of catalyzing scientific breakthroughs and accelerating new discoveries in ways previously unimagined. EarthCube aims to infuse emerging practices and technological innovations in digital scholarship, data science and analytics, data and software stewardship, and open science, into basic research in the geosciences. EarthCube not only fosters and leverages advances in data and information science and technology, but also brings together new communities of practice that engage in cyberinfrastructure related activities.

EarthCube's distinctive value is in its vision to address technical, organizational, social, and cultural issues in an **integrative** and transdisciplinary manner. Critical goals for EarthCube are:

- integration across scientific communities that share common cyberinfrastructure needs and challenges in the geosciences;
- integration of geoscience communities with computer, information, and engineering science communities;
- integration of multi-disciplinary data, models, and software tools;
- integration of existing cyberinfrastructure resources, capabilities, and expertise with new technologies and emerging practices of digital scholarship

Integration is EarthCube's objective and imperative for achieving its strategic vision of empowering scientists at the frontiers and enabling a transformation in geoscience research.

This document outlines a broad roadmap, including a series of action plans divided into near-, mid-, and long-term time frames designed to leverage the suite of existing activities and provide an integrated path toward realizing the goals outlined in the EarthCube Strategic Vision Document. The EarthCube Roadmap describes a set of priority activities to fulfill the promise of EarthCube's diverse components including scientific and technological objectives, educational aspirations, community organization and engagement principles, and governance structure. The EarthCube Leadership Council has identified these goals and action plans based on recommendations from the EarthCube community, including the governance teams and committees and working groups.

A. EarthCube Goals[1]

1. Science Goals

The primary scientific goals of EarthCube are to enable significant progress in our understanding of fundamental Earth system processes, and to improve our ability to mitigate complex, large scale environmental problems, by fostering innovations in data science and information technology. Accelerated scientific progress in such topics necessitates integrated and data-intensive domain-specific studies at ever increasing spatial and temporal resolution, as well as multidisciplinary interactions across different science domains. EarthCube-enabled science relies fundamentally on open access to well-documented data[2], models, software, analytical and visualization methodologies and tools. The motivation for EarthCube-enabled science is underpinned by three fundamental goals[3]:

- To characterize the key processes, interactions, causations, and feedbacks operating at and across different temporal and spatial scales within physical, chemical, and biological domains.
- To quantify limits of prediction and better understand the constraints on and limits of data and model accuracy and utility.
- To deliver a holistic, quantitative representation of critical Earth physical, chemical, and biological states and fluxes, in order to inform fundamental science and societal decisions.

2. Technology Goals

EarthCube technology and related cyberinfrastructure will function as a bridge between data and knowledge. In so doing EarthCube will accelerate geoscientists' ability to characterize and understand complex Earth systems, by providing enhanced access to data, and new technologies and methods to integrate, analyze and visualize those data. Addressing EarthCube science goals necessitates the development of sophisticated and rigorous data-enabled scientific inference methods, i.e., technologies for learning from complex and ever-increasing data volumes, as well as methods that will guide what/how data should be obtained for effective scientific analysis and understanding. As such the core tenets of EarthCube technology are:

- Knowledge-rich environments: Scientific resources and products will be described with rich metadata to enable understanding and reuse. Advanced analysis and visualization

[1] The EarthCube process to date includes a volunteer, community-driven committee structure, i.e. the EarthCube Test Governance, the EarthCube funded projects, as well as engagement from the wider geoscience community. Under the leadership of the Science Committee (SC) and the Technology and Architecture Committee (TAC), a science plan and a technology plan were developed. At the time of writing, these plans are draft documents in the process of community review as presented during the 2015 EarthCube All Hands Meeting. These plans have been synthesized and further refined by the EarthCube Leadership Council (LC) into the draft EarthCube Strategic Vision document which contains a set of overarching science and technology goals summarized below to provide context for the roadmap.

[2] Here data broadly include observational and interpretive, model, experimental, and physical sample items.

[3] See EarthCube's Strategic Science Plan – Aronson E.L. et al., 2015, Geoscience 2020: Cyberinfrastructure to reveal the past, comprehend the present, and envision the future, EarthCube Working Paper, ECWP-2015-1, 19 p. http://dx.doi.org/10.7269/P3MG7MDZ

techniques will be developed to enable integration and automated learning techniques from data to take advantage of knowledge-rich components for effective mining, fusion, and assimilation of data into sophisticated inference models.

- Open science: EarthCube will promote the design and creation of discoverable and accessible scientific resources and products, in order to foster cross-domain dissemination and integration, enable reproducibility, and ensure that they can be adapted to solve new problems.
- Federated organization: EarthCube participants, from organizations to individuals, will contribute resources designed to interoperate through agreed standards and protocols. Rather than centralized control, EarthCube will provide coordination by fostering standards and integration in a loosely-coupled distributed organization.

B. EarthCube Principles

The vision for EarthCube is based on a set of core principles that guide and underpin the roadmap for EarthCube's development, implementation, and operation. EarthCube Governance and decision makers must continuously balance competing views and but be grounded by the guiding principles that are the overarching, exemplary concepts and commitments. These principles are also recognized as key elements for success:

- Maintain focus on community, innovation, and ease of use.
- Foster science-driven and service-oriented culture across all areas.
- Advocate for Open Source software, tools, and technologies.
- Adapt, adopt, and only when necessary, develop new tools and technologies.
- Emphasize sustainable and modular solutions that are scalable and interoperable.
- Harness, build on, and integrate capabilities developed through other investments, encouraging federation of distributed systems.
- Cultivate and proactively leverage collaborations and partnerships.
- Improve continuously and be open to new and divergent perspectives.
- Employ iterative and agile development model for all elements.
- Promote self-assessment and accountability and develop metrics to monitor progress.
- Communicate with consistency and clarity on a continuous basis.

C. EarthCube Organization

1. EarthCube Community

At the heart of EarthCube is an inclusive, evolving, dynamic community of researchers, students, engineers, data managers, and other professionals, as well as partner organizations from across the geosciences and cyberinfrastructure sectors that work together to guide, govern, and advance EarthCube.

Currently, the EarthCube community encompasses more than 2,500 members with broad scientific and engineering interests and expertise. Participation ranges across a wide spectrum of activities from involvement in funded projects, governance committees, or working groups to organization or attendance at EarthCube workshops and meetings, to subscribing to the EarthCube newsletter. Mechanisms that have been particularly effective in growing membership have been EarthCube's domain end-user workshops and the funded Research Coordination Networks, which build and engage communities through workshops, field trips, forums, and webinars. These efforts have lead to self-organized working groups and, most importantly, to new collaborations and projects.

EarthCube's Engagement Team (ET) and Liaison Team (LT) both focus on the growth of the EarthCube community. These teams carry responsibilities for defining strategies and identifying opportunities to ensure community engagement (ET) and for linking the activities of the wider-EarthCube effort to relevant organizations and initiatives globally (LT). In turn, the EarthCube Project Office carries critical responsibilities toward the community, including development and maintenance of the EarthCube web site for information dissemination and collaboration workspaces; regular communications such as weekly updates and monthly newsletters; organization of the annual All-Hands Meeting; organization of exhibit booths, town halls, and other events at geoscience conferences; support of governance committees; and creation of brochures and other promotional materials.

2. EarthCube Governance

The EarthCube Test Enterprise Governance Project (ECTEG) is an ongoing experiment to test community-led recommendations for a governance framework aiming to further the National Science Foundation's EarthCube goals of developing operational cyberinfrastructure to support domain scientists and other end-users in the geosciences. Over a two-year period, this project:

- Selected an appropriate and community-agreed governance framework;
- Vetted this EarthCube Enterprise Governance framework with the community and the National Science Foundation;
- Demonstrated and evaluated the specific charter, by-laws and terms of reference of the governance framework in a pilot; and
- Proposed next steps for implementing longer-term EarthCube governance.

In the first year of this project (Stage I - Planning), the project carried out extensive community engagement across multiple levels and venues, which led to a community-developed organizational framework for EarthCube. In the second year (Stage II - Demonstration), the project implemented this organizational framework, and tested the experiment in multiple ways to make progress towards the goal of developing a longer term enterprise governance framework for EarthCube. A number of elements were used to test and stress the Demonstration Governance Pilot including selection and election of leaders; writing charters for committees, teams, and the Leadership Council; engaging and communicating with EarthCube stakeholders, particularly end-user geoscientists; and engaging and incorporating the NSF-funded EarthCube projects, including integration of the Conceptual Design award outcomes.[4]

D. State of EarthCube Technology R&D

EarthCube technology research and development has thus far taken place mostly in the context of funded projects. These funded projects have so far included Building Blocks (15), Conceptual Designs (3), Research Coordination Networks (7), and Integrative Activities (13). Building Block activities develop technologies designed to address the needs of EarthCube geoscientists

[4] Other Test Governance activities included: Setting up a budget process; Developing the roles of the Program Office Manager POM vis-à-vis the governing Leadership Council, NSF, and the NSF-funded Program Office; negotiating a services plan between the office and the Leadership Council as a model for a proposed longer-term enterprise governing body; preparing a science vision for EarthCube; considering funding of volunteers in leadership roles; continually and interactively developing the organizational framework to address NSF requirements; Involving end-users; and creating an EarthCube community of practice.

including such things as data brokering and interoperability, enhanced metadata profiles, real-time data streams, ontology design patterns, and semantic infrastructure. The conceptual design efforts surveyed the various types of approaches to developing an EarthCube architecture. These efforts are synthesized by the TAC Architecture Working Group. Research Coordination Networks are important elements of community building, developing use cases, and defining science and cyberinfrastructure needs. Integrative Activities are the most recent set of projects to be funded and the goal is to extend community engagement in the development of standards and data infrastructure.

A preliminary analysis of the distribution of these various activities with respect to geoscience domains and type of technology design, development, or assessment, reveals that both the cyberinfrastructure needs, as well as the stages of development in the adoption of data science and technology into geosciences, vary considerably in scale, scope, and characteristics across the different geoscience disciplines. As such, it is important that the prioritized actions for EarthCube research, development, and education be broad, agile, and allow sufficient flexibility to accommodate the varying needs of all geosciences disciplines.

II. Realizing the EarthCube Strategic Vision - The Roadmap

EarthCube development is, by design, phased. The first phase (2011-2013) was a conceptualization phase, phase two (2014-2015) focused on testing and refining the developing community processes, and phase three (2016-) is focused on implementation.

The integrative vision of EarthCube requires that technology development activities address not only traditional cyberinfrastructure constructs such as repository development or metadata capture, but also technological innovations in the emerging practices of digital scholarship, data science and analytics, data and software stewardship, and open science. The technology R&D strategy must strive to make disciplinary boundaries permeable, nurture and facilitate knowledge sharing, cultivate unanticipated uses of information, and enhance collaborative pursuit of cross--disciplinary research. It must address the paradigm of open science and open data in a holistic way.

Stewardship of contributed data (including data rescue and the liberation of dark data), software, and research products through open science practices will be the foundation for all EarthCube-enabled science. This stewardship allows access to the high quality, interdisciplinary data necessary to characterize key Earth processes across relevant spatial and temporal scales. Knowledge-rich technical components will facilitate data mining and advanced data analysis, through computational, statistical and algorithmic techniques in order to quantify limits of scientific prediction, and constraints of modeling efforts. Distributed and loosely-federated technologies using standard services will enable remote access, visualization and interoperability of complex distributed resources, allowing synthesis of cross-disciplinary knowledge and information sharing to inform new science and societal decisions.

The remainder of this EarthCube Roadmap is organized along a set of themes that collectively implement the EarthCube strategic vision. These themes form the pillars upon which EarthCube will be developed and against which EarthCube can be evaluated. These themes include Infrastructure, Education and Capacity-Building, Assessment, Community and Communications, Governance, Collaborations and Partnerships, and Investments and Sustainability.

A. Infrastructure

Thus far, the process of assessing the current state and identification of future infrastructure needs has been primarily undertaken by the Technology and Architecture Committee (TAC) with interaction with the EarthCube Leadership Council. The activities of the TAC are currently organized based on its original charter, and include: 1) organize the collection of science use cases to derive technology requirements, 2) analyze the funded EarthCube projects to detect gaps, 3) compile initial requirements for a testbed, 4) initiate documentation and coordination for standards activities, 5) design a community roadmap for EarthCube architecture. The TAC engaged the community through Working Groups to pursue these activities, and conducted a comprehensive survey of EarthCube funded projects that covered topics of interest to all these Working Groups[5]. The working groups generated an initial Technology Roadmap[6] and subsequent recommendations were provided to the Leadership Council as input. The Leadership Council used these inputs as well as other previous EarthCube activities, such as the end user domain workshops to create the synthesis provided here.

1. Architecture

Goal: Provide a high level description of the EarthCube system architecture, required capabilities, and a set of structures to serve the data and information needs of the geoscience community in a sustainable manner.

Rationale: Up to the current phase of EarthCube, most conceptual perspectives on the nature and role of EarthCube architecture have emerged from the EarthCube Conceptual Design funded projects. The TAC Architecture Working Group, engaging the community through discussions and workshops, has explored the set of emerging principles to converge on the design of EarthCube architecture.[7] To facilitate integrative science, key considerations of the architecture include, but are not limited to, standards adoption, semantic technologies, cloud infrastructure, and shared high-performance computing resources.

Near-term Actions:

- Characterize current architecture, based on existing components and infrastructure composed of different systems, facilities, and technologies.
- Crystalize the definition of EarthCube architecture based on EarthCube strategic vision.
- Engage the science and cyberinfrastructure communities to converge on a consensus of a description of EarthCube architecture and what it must entail.

Mid-term Actions:

- Leverage the EarthCube testbed and other test environments for capabilities to verify that integration, standards and interoperability requirements are satisfied.

---

[5] The survey results are included in the TAC Gap Analysis Working Group report: https://docs.google.com/document/d/1rhjRPJLukGmKwzYqgr4X-DoX_iTX2GyX9IfRh0mYuDY/edit?pref=2&pli=1 and summarized in this publication: Smith II, P. L., Malik, T., & Berg-Cross, G. (2015). Rediscovering EarthCube: Collaborate. Or collaborate not. There is no I. OCLC Systems & Services: International Digital Library Perspectives (OSS:IDLP), v. 32(3).

[6] The initial EarthCube Strategic Technology Plan, released at the EarthCube All Hands Meeting in May 2015, is available here: http://earthcube.org/document/2015/strategic-technology-plan.

[7] The full Architecture WG Final Report is available here: https://docs.google.com/document/d/1X7f17c44aRutHaxKfpevurzuywh95Sg6LtiP_pEIlIo/edit?usp=sharing. The Architecture WG's proposed Roadmap is here: https://docs.google.com/document/d/1phzixOYhwrTK9Viuy_KpsMeUIk30LBnnCzOntVsMGdk/edit?usp=sharing.

- Establish an organizational role and a process for the selection of an architecture committee (and possibly a chief architect) with the role of maintaining architectural knowledge base.
- Engage international communities, such as GEO and Digital Earth, in the architecture definition and development process.

Long-term Actions:

- Systematically and periodically conduct reviews of the architecture considering advancements in science and technology, additional prototypical use cases, and system monitoring.
- Develop a mechanism to sustain EarthCube architecture for continuously channeling community input into architecture development, through geoscience-focused workshops, demonstrations, conference sessions, telecons, online live documenting, and other activities.
- Maintain a living architecture document for long term evolution of EarthCube.
- Engage diverse experts to evaluate, verify and validate EarthCube's architecture.

2. Use Cases: Utility, Collection and Synthesis

Goal: Develop a comprehensive knowledge-base of geosciences-related cyberinfrastructure needs.

Rationale: A systematic approach to identify the diverse cyberinfrastructure needs of the geoscience community is essential for guiding the development of EarthCube's architecture. One such approach, currently undertaken by the TAC Use Cases Working Group utilizes documented descriptions of geoscience research objectives and the corresponding sequence of steps taken to achieve them[8]. Collection and synthesis of scientific use cases, if sampled in a comprehensive and representative fashion, facilitate user-driven development of EarthCube by revealing computational and data gaps in the current cyberinfrastructure, thereby guiding development of an open, knowledge-rich, federated system necessary to support EarthCube-enabled science. Use cases facilitate communication and promote collaboration among geoscientists and cyber-related scientists, leading to stronger engagement across the EarthCube community, and novel discipline-specific and cross-disciplinary research.

Near-term Actions:

- Assess the effectiveness of use cases in obtaining a comprehensive and representative characterization of geoscience data and cyberinfrastructure needs.
- Complete compilation of geoscience use cases through interviews with geoscientists, and literature review of existing EarthCube resources; determine if additional collection is warranted, and if so, what are the appropriate sampling strategies to ensure comprehensive coverage.
- Create a use case repository that will function as a shared resource for both science and technology objectives.

Mid-term Actions:

- Further refine the repository structure facilitated through direct engagement of the EarthCube science and technology communities.
- Synthesize requirements expressed in use cases, including identification of frequently recurring needs.

Long-term Actions:

[8] The Use Cases Working Group report and ongoing collection are described here: http://earthcube.org/group/use-cases-wg

- Develop strategies for sustained and adaptive determination of requirements to ensure that EarthCube architecture, tools and services are responsive to evolving science needs.

3. Testbed

Goal: Capture products of funded projects, allowing scientists to easily explore and experience new technologies, and enable experimentation with the integration of different technology components.

Rationale: An effective testbed environment would: 1) Serve as an infrastructure for test and evaluation of software and EarthCube components, capabilities, and services; 2) Foster integration and interoperability across funded projects, data facilities, as well as with other community infrastructures; 3) Preserve the outcome of funded projects. The TAC Testbed Working Group has developed requirements for an EarthCube testbed by creating an inventory of needs and requirements[9].

Near-term Actions:

- Rapidly implement and demonstrate a prototype of the EarthCube testbed infrastructure and interfaces, based on a simple use case.
- Develop mechanisms for community engagement and participation to determine whether the EarthCube testbed is addressing major needs.

Mid-term Actions:

- Implement and document best practices for evaluation of future technologies.
- Identify and recommend ways in which the EarthCube testbed could be used to demonstrate and evaluate the emerging EarthCube architecture.

Long-term Actions:

- Provide an operational EarthCube testbed environment to perform ongoing evaluations of the results of funded projects as well as other technologies.

4. Standards

Goal: Promote adoption and use of data, metadata, and related information technology standards to facilitate integration and interoperability of data and cyberinfrastructures.

Rationale: Standards are essential to enable reusable EarthCube cyberinfrastructure, data exchange and interoperability, and to support provenance and scientific reproducibility. EarthCube needs a framework for recognition and adoption of relevant technology standards, and for possible coordination with international standards organizations. The TAC Standards Working Group has developed an initial compilation of the standards being used in current EarthCube funded projects, and produced a set of recommendations to foster the adoption of standards[10].

---

[9] The report from the Testbed WG is available here: https://docs.google.com/document/d/1GlvrQ1A4dwt0LWTivaxA7tu5pWnkPjo77vPFA9vy7Q/edit?pref=2&pli=1#heading=h.iec5kdnkr5m1.

[10] The report from the Standards WG is available here: http://earthcube.org/document/2015/ecstandardsrecs

Near-term Actions:

- Develop and maintain a community registry of technology standards used in EarthCube projects and products.
- Develop processes and workflows for coordinating standards activities.
- Develop, utilize and promote an EarthCube glossary for consistent use of terms across the EarthCube community.

Mid-term Actions:

- Create an EarthCube Standards Authority for overseeing the standards processes, resolving competing or conflicting use of standards, and organizing coordination with external standards bodies.
- The Liaison Team will work to identify opportunities for collaboration with external standards bodies and other organizations on relevant standards development activities, e.g., RDA, OGC, W3C, etc.

Long-term Actions:

- Collect community feedback on standards to report on adoption and to collect requirements for further standardization needs.

5. Practices and Policies

Goal: Promote sharing, adoption and use of community best practices in data management and technology development; inform and adhere to existing policies of agency, organizational and institutional stakeholders.

Rationale: EarthCube operates in an environment of federal, organizational, and institutional policies. It is essential to align data management and technology development activities with existing relevant stakeholder policies and community best practices to facilitate interoperability of information and create a federated cyberinfrastructure system. Therefore, EarthCube needs a framework for recognition and adherence of relevant policies and best practices. In particular, data related developments will promote the following data principles:

- Enhanced access to data
- Rich metadata
- Discoverable, accessible data
- Enable reproducibility of data
- Open science
- Data quality
- Open Source
- Scalable
- Interoperable
- Adaptation for reuse
- Foster common standards
- Sustainable

Near-term Actions:

- Develop strategies to engage the Council of Data Facilities (CDF) to leverage existing and emerging policies and best practices with respect to data management and related activities.
- Develop a plan to identify, inform and promote compliance of relevant technology policies within the EarthCube community.
- Define who in the EarthCube governance will manage and disseminate information regarding stakeholder policies and community best practices.

- Develop these plans in coordination with the findings and actions from the Standards WG, outlined above.

Mid-term Actions:

- Recommend guidelines (and requirements) for data management practices such as Data Management Planning.

Long-term Actions:

- Assess compliance and develop strategies to reinforce adherence to relevant policies.
- Continually review and update EarthCube's practices and policies.

B. Education and Capacity Building

Goal: Establish education, training, and workforce development activities that are essential to realizing the EarthCube vision of transforming the conduct and culture of geosciences.

Rationale:
The fundamental role of education and training of the next generation of geoscientists and information technology professionals in modern data-intensive science is pivotal to realizing EarthCube's strategic goals and sustaining its long-term success. The intrinsically intertwined nature of scientific pursuit and human capacity-building necessitates strategic activities that will advance data and information science literacy of students and early-career geoscientists.

Near-term Actions:

- Promote increased interaction and closer collaboration among geoscientists, data and information scientists, and educators with the goal of identifying the necessary skills for data-intensive scientific research.
- Define and articulate the essential role of interdisciplinary educational activities to realize the integrative goals of EarthCube.
- Explore the integration of research and data in pedagogy and cross-disciplinary curriculum development into the geoscience programs.

Mid-term Actions:

- Highlight career paths at the interface of geosciences and data sciences.
- Promote establishment of cross-disciplinary educational programs into traditional geoscience curricula.
- Identify best venues to share EarthCube products with educators.

Long-term Actions:

- Ensure that EarthCube products will become valuable tools in geoscience educational activities and curricula.
- Promote pathways to support the development of career paths of a cadre of technical experts within geosciences.
- Explore possibilities for providing professional recognition for data-related expertise in geosciences, e.g., preservation, curation, and dissemination.

C. Assessment

Goal: Analyze and optimize the value of EarthCube investments through the application of a comprehensive assessment framework.

Rationale: A comprehensive assessment framework able to address process, input, output and impact metrics is needed for EarthCube. Process metrics will measure the effectiveness of the governance and specific actions taken to achieve the EarthCube Strategic Vision. Input metrics will assess quantifiable inputs into efforts to achieve the Strategic Vision. Output metrics will measure the tangible outcomes and products from EarthCube. Finally, impact metrics will assess the overall success of EarthCube activities and products in advancing EarthCube science goals.

Near-term Actions:

- Develop a comprehensive assessment framework through stakeholder participation. Consider such elements as:
    - Establish an assessment organizational structure (TBD committee, team, working group) that includes representatives of diverse stakeholders,
    - Gather community requirements for assessment, and
    - Hold a workshop to begin design of the framework.
- Develop appropriate quantifiable metrics for each assessment task.

Mid-term Actions:

- Apply the assessment framework including continual monitoring of metrics.
- Provide assessment results in a publicly accessible venue.
- Evaluate the effectiveness of the assessment framework in collaboration with stakeholders and make necessary improvements.

Long-term Actions:

- Provide assessment results in a publicly accessible venue.
- Monitor and communicate impact metrics.

1. Gap Analysis

Goal: Identify gaps in EarthCube technological capabilities and investment strategies as determined by community requirements.

Rationale: There is a need to aggregate, coordinate, and articulate accurate and comprehensive information about the EarthCube funded projects and existing capabilities within the geosciences, and to identify gaps in capabilities, investments, and developments. The gap analysis will help inform the future development of EarthCube architecture.

Near-term Actions:

- The TAC Gap Analysis Working Group has initiated an assessment of the technology areas under development by the EarthCube funded projects. The effectiveness of this approach must be assessed and insights from the assessment should guide its course toward completion.
- The Use Case Working Group has been collecting use cases to help identify community requirements and current gaps in capabilities. The effectiveness of this approach must be assessed, in particular verifying that the use case collection covers a representative sample of all geosciences; insights from the assessment should guide its course toward completion.
- Establish a common repository to collect and communicate the results of gap analysis.

Mid-term Actions:

- Use results of gap analysis to provide guidance to stakeholders on EarthCube architecture development.

Long-term Actions:

- Perform continual/periodic reassessment of the approach taken to gap analysis and refine according to the changing landscape of geosciences and related technology.

D. Community and Communications

Goal: Ensure the widest possible participation of relevant stakeholders in EarthCube and work to ensure the widest possible use of EarthCube.

Rationale: As a community-driven and led effort, EarthCube's success depends on the ongoing engagement of a diverse community of stakeholders. Building and maintaining trust in the EarthCube process and the quality and utility of its products is essential and needs careful attention and dedicated resources both in the near and in the long term. EarthCube needs effective strategies aimed at addressing the "3C" challenges of clarification, communication and collaboration in order to engage individual geoscientists across the EarthCubecommunity, but also in order to foster networking, collaboration, alignment, and integration of the different entities of the EarthCube governance including the EarthCube Project Office and the NSF, the funded EarthCube projects, and related initiatives and organizations relevant to EarthCube's mission.

Near-term Actions:

- Articulate benefits to members and create a communications strategy and plan, including evaluation of strengths and weaknesses of current engagement programs and their effectiveness in reaching the target communities.
- Define metrics to assess community involvement.
- Ensure that the scope of community engagement is broad expanding appropriately beyond geosciences and includes communities is data and information science and engineering with the relevant expertise for EarthCube.
- Improve document management (including formal publication of workshop reports, technical reports, etc.), and communication about consistent usage among committees and teams.
- Identify needed improvements to the EarthCube web site that enhance access to information, including how to engage in EarthCube and about EarthCube products and developments.
- Identify strategies to increase the engagement of funded projects.
- Promote volunteer recognition.
- Evaluate integration of EarthCubeWiki with main EarthCubewebsite, and make a recommendation.

Mid-term Actions:

- Develop mechanisms to enable engagement and support of individual end users and stakeholders (e.g., professional societies, publishers, government, commercial), including attracting new users to EarthCube (collaborating with the EarthCube Liaison Team where appropriate).

Long-term Actions:

- Design and develop structure to support all of the outreach strategies (outlined above), with an emphasis on sustained stakeholder leadership.

E. Governance

Goal: Provide, assess, and evolve an organizational structure to lead and manage EarthCube.

Rationale: EarthCube Governance is the structure which helps to develop and translate community consensus into tangible results. Governance helps to provide ongoing continuity and decision-making to guide EarthCube in achieving its vision. Finding the correct balance between the overhead that governance places on process and community-led outcomes is an ongoing exploration implicit in the governance activity.

Near-term Actions:

- Work in concert with the assessment process to develop appropriate metrics and apply them to evaluate the effectiveness and legitimacy of current governance structure.
- Finalize, adopt, and implement the EarthCube charter as a necessary step for EarthCube enterprise governance.
- Engage the community in identifying ways to improve the transparency of decision-making processes within the governance structure.
- Establish effective pathways between the EarthCube Enterprise Governance and NSF for communicating community needs and recommendations in advancing EarthCube's scientific, intellectual, and technological goals.
- Clarify the role, and finalize the services plan that outlines the responsibilities of the proposed Science Support Office to the governance organization and with NSF.
- Address the functional challenges in the current governance structure with respect to transparency to increase trust in decisions and lead to greater legitimacy for the governance structure of EarthCube.
- Re-examine whether, under what terms, and which leaders will receive financial compensation, and if so, by what mechanisms; and consider the effect of any financial compensation on the community-based nature of EarthCube.

Mid-Term Actions:

- Designate an organizational host to replace the temporary Project Office to maintain continuity and momentum in EarthCube's community governance structure and activities and serve as trustee of NSF (and other) funds allocated to support EarthCube governance and organization.
- Structure funding mechanisms and organizational requirements to ensure that the EarthCube cyberinfrastructure is driven by, and meets the needs of, the geoscience research community and its end-users.

Long-term Actions:

- Continually assess the evolving structure of the governance and its effectiveness.

F. Collaborations and Partnerships

Goal: Develop partnerships and collaborative activities with aligned entities and organizations to advance achievement of the EarthCube Strategic Vision.

Rationale: Strategic partnerships are crucial to the success of a community-driven and resource-constrained effort like EarthCube. The "System of Systems" thinking put forth by the EarthCube Architecture Working Group requires that EarthCube leverages the resources, capabilities, human capital, and expertise residing in other organizations engaged in related activities. By fostering partnerships, EarthCube can benefit from the work done in other organizations, conduct mutually beneficial workshops and meetings, and pool resources for education and outreach activities to increase our understanding of earth systems and progress toward the EarthCube Strategic Vision.

The Engagement Team and Liaison Team in EarthCube's governance are expected to play a significant role in cultivating such partnerships by identifying potential collaborations, leading to the entrainment of innovative tools and data assets of importance to Earth System science research and education, advancing standards, and increasing interoperability of tools and services across the geosciences landscape.

Near-term Actions:
- Identify key prospective organizations for collaboration.
- Develop documentation to formalize collaborations and partnerships.
- Establish policies and procedures for developing formal relationships.

Mid-term Actions:
- Establish collaborations and partnerships to advance EarthCube Strategic Vision, including entities within NSF and other funding agencies, professional organizations, and international partners.
- Monitor collaborations and partnerships to ensure mutual benefit and productivity.

Long-term Actions:
- Actively maintain and establish new partnerships and collaborations.
- Broadly communicate the impact and outcomes of these partnerships.

G. Investments and Sustainability

Goal: Ensure the long-term reuse and sustainability of EarthCube resources and capabilities to maximize the return on investment and utility for science.

Rationale: Moving forward, development of a research-to-operations workflow, infrastructure, and support will be critical to maximizing the return on investment of NSF investments. The initial EarthCube program funded a number of projects with exciting R&D outcomes. How will outcomes be made available and supported going forward? This brings up the forward-looking challenges of evaluation, iteration, and ongoing operations of selected EarthCube products.

Near-term Actions:
- Establish an assessment framework to evaluate products that need to be sustained.
- Use the results of this assessment to guide the sustainability of those products, including the design and scope of requirements of future solicitations.
- Identify and recommend new mechanisms for investments in EarthCube research and development that consider the important factors of comprehensive technical coverage, plans for sustainability of outcomes/products, close cross-disciplinary interaction, and related educational activities.
- Establish partnerships with organizations that may provide infrastructure to support EarthCube product sustainability e.g. federally funded data facilities through the Council for Data Facilities.

Mid-term Actions:
- Provide guidance to stakeholders on potential support modes for sustainment of EarthCube products and services.
- Monitor and evaluate results of EarthCube-like activities in other regions, such as the international Belmont Forum and The Research Data Alliance among others.

Long-term Actions:
- Coordinate with NSF and the EarthCube Advisory Committee for continued evaluation of progress, and potential directions for growth and adaptation.