REPORT: EARTHCUBE ALL-HANDS MEETING 2014

June 24–26, 2014
Washington, DC

Prepared by the EarthCube Test Enterprise Governance Team, August 2014

EARTHCUBE 2014 ALL-HANDS MEETING

CONTENTS

This material is based upon work supported by the National Science Foundation under Grants No. 1340233 and 1417948. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

EXECUTIVE SUMMARY

The 2014 EarthCube All-Hands Meeting was held June 24–26, 2014 in Washington, D.C. Nearly 150 members of the geo- and cyber-science communities, representing over 90 academic institutions, government agencies, non- and for-profit organizations, and consortia were brought together to share their progress and experience with EarthCube thus far, and discuss and plan activities for the coming year. The Meeting had three principal goals:

1. **Come to convergence on an EarthCube Demonstration Governance Charter for use during a pilot phase (October 2014–September 2015)**
2. **Examine the state of cyberinfrastructure in the Geosciences**
3. **Provide a forum for communication among the Geosciences and technology communities**

Meeting goals were met through a mix of both facilitated and non-facilitated plenary and breakout sessions or business meetings as well as impromptu sessions, group meetings, and ad hoc discussions/presentations. Over 25 session proposals brought forth by the community were included in the agenda to help create a dynamic and diverse community-driven meeting. By achieving the meeting goals, a series of outcomes and recommendations (related to governance, technology and architecture, and community building) came out of the All-Hands Meeting, positioning the EarthCube community to enter its Demonstration Governance phase.

Governance Outcomes

The primary governance-related goal of the All-Hands Meeting was to facilitate convergence on the EarthCube Test Enterprise Governance Charter to be implemented during the demonstration period (October 2014-September 2015). This goal was carried out via a series of plenary and breakout sessions designed to gather participant feedback on the Charter, and to allow for real-time, participant-driven changes. The first governance session resulted in insightful observations and suggestions on refining the proposed governance structure, summarized here:

- The overall goals, organization, and structure of the draft governing system are representative of the community's needs and are appropriate for moving EarthCube forward, but additional work is needed to flesh out the initial design.
- The community has a strong interest in further defining the scope of EarthCube and EarthCube governance to guide effective decision-making and community interaction.
- More work is needed to clarify roles and responsibilities within the governance components, as well as the interactions between those components.
- Decision-making processes and lines of authority will also need additional consideration over the coming demonstration year.

Subsequent sessions focused on addressing these observations and other outstanding issues, resulting in a well-defined governance structure (Figure 1) and a set of priorities and actions to address regarding the implementation of that structure. The emergent governance structure components and recommendations are as follows:

- **Leadership Council**: Determines the scope/vision and maintains the strategic direction of EarthCube; acts as single communication point with NSF.
- **Science Standing Committee**: Determines policy/strategy on data management plans; engages scientific community (along with Engagement Team).
- **Technology and Architecture Standing Committee**: Facilitates community convergence on a system architecture (in partnership with Science Committee and Council of Data Facilities); coordinates a testbed for development of EarthCube components.
- **Council of Data Facilities Standing Committee**: Facilitates shared services through data providers.
- **Engagement Team**: Develops a structured communications plan; compiles and disseminates science end-user use cases.
- **Liaison Team**: Initiates and develops implementation of a strategy for collaboration with related activities across agencies, industry, international, and other science domains.
- **Science Funded Team/Technology Funded Team**: Act as coordinating mechanisms across EarthCube funded projects.
- **Working Groups:** Ad hoc but sanctioned organizational units created in response to a pressing issue, opportunity, activity, or deliverable related to EarthCube's goals.
- **Special Interest Groups:** Informal ad hoc groups that self-organize around areas of common interest. They may evolve into formal Working Groups.
- **Membership:** During the demonstration governance phase, membership will be free and open at an individual and institutional level in order to be as inclusive as possible.

Technology and Architecture Outcomes

A two-day Architecture Forum, organized by members of the Test Enterprise Governance IT Team, convened at the All-Hands Meeting with the primary goal of accelerating community convergence on an EarthCube architecture. A secondary goal of the workshop was to discuss the charter of the Technology and Architecture Standing Committee, to be put into place October 1, 2014 as part of the emerging Demonstration Governance structure, as well as to refine the Committee's functions and responsibilities as presented in the Test Governance Charter document.

The group concluded that concrete decisions about the architecture of EarthCube cannot be made until the governance organization is in place; however, work is underway to seek consensus on the nature of the required architecture specifications and how they will be used. Several points of general agreement emerged from the Architecture Forum plenary and breakout sessions, resulting in a set of organizational and design recommendations to inform activities of the Technology and Architecture Standing Committee, tasked with the stewardship and development of an EarthCube architecture.
A series of facilitated large- and small-group sessions resulted in the following key accomplishments:

- Made significant progress toward a shared understanding of what constitutes an ongoing architecture.
- Identified critical issues of concern to community members regarding an EarthCube architecture.
- Assembled representatives from all of the NSF-funded EarthCube Building Block awards to determine potential alignments, interfaces, and architecture needs, establishing an inventory of emerging architecture elements based on the EarthCube funded projects.

- Collaboratively revised elements of the Technology and Architecture Committee Functions as outlined in the EarthCube Test Enterprise Governance Charter.
- Through the use of a collaborative "living glossary,"[1] created a basis for future communication across the EarthCube technology community.

Community Building Outcomes

Outcomes of the All-Hands Meeting included several recommendations for next steps related to engagement and community building, in order to successfully move into and carry out the governance Demonstration Phase. The first of these was to inform the larger EarthCube community of the results of the All-Hands Meeting and how to get involved in the project moving forward. The following timeline was disseminated via the EarthCube Newsletter, social media, listservs, and email outreach:

- July 17, July 21, and Aug. 5: Informational webinar on the new Standing Committees and Teams
- Late August: First Standing Committee and Team virtual meetings
- August 31: Deadline for nominations of At-Large Leadership Council members
- September 15: Deadline for selection of At-Large Leadership Council members

Meeting participants also identified a need to launch a renewed community outreach program aimed at populating the demonstration governance structures (*Science Standing Committee, Technology and Architecture Standing Committee, Engagement Team, and Liaison Team)* in order to effectively test the proposed Charter and allow for elements to be adjusted to better meet community needs. There was also a call for greater engagement and collaboration with and among the EarthCube portfolio of funded projects.

[1] Glossary available at: http://bit.ly/1q8NWqe

Introduction and Background

The 2014 EarthCube All-Hands Meeting was held June 24–26, 2014 in Washington, D.C. This meeting represented the culmination of a 10-month community engagement process (September 2013–June 2014) and was intended to highlight the work and recommendations of the EarthCube community, including members of funded projects as well as the broader array of stakeholders and volunteers. The Meeting had three principal goals, the outcomes of which were intended to position the EarthCube community to enter a demonstration governance phase. These goals were:

1. Come to convergence on an EarthCube Demonstration Governance Charter for use during a pilot phase (October 2014–September 2015).
2. Examine the state of cyberinfrastructure in the Geosciences.
3. Provide a forum for communication among the Geosciences and technology communities.

The first goal was accomplished through facilitated plenary and small group discussions, as well as through interactive scenario exercises designed to test the proposed governance structure. These sessions resulted in consensus on revisions to the Charter and governance structure, as well as a set of priorities to be addressed by this structure during the demonstration year. The second goal was accomplished via the sharing of results of existing projects, including EarthCube award projects and other projects and initiatives not funded by EarthCube grants. This was done during the EarthCube Architecture Forum, as well as through many of the community-proposed sessions. The third goal was reached via numerous networking opportunities and community-proposed sessions, and will extend into the coming year and beyond with community participation in the organizational units formed under the proposed governance structure.

This report is intended to serve as a public record documenting the outcomes, processes, and rationales of this meeting related to governance, technology and architecture, and community engagement. It is not intended to reflect any changes in community and organizational direction that may take place in the months following the meeting.

Outcomes and Process

Governance Outcomes and Recommendations

Demonstration Governance Structure

The primary governance-related goal of the All-Hands Meeting was to facilitate convergence on the EarthCube Test Enterprise Governance Charter to be implemented during the demonstration governance period (October 2014–September 2015), and to poise EarthCube to move forward into the 2014–2015 demonstration governance phase. This goal was carried out via a series of plenary and breakout sessions designed to gather participant feedback on the Charter, and to allow for real-time, participant-driven changes. These sessions resulted in a well-defined governance structure and a set of priorities and actions to address regarding the implementation of that structure.

On the morning of day three of the meeting, the following changes to the Charter were presented in the closing plenary session (other elements of the charter that had no significant changes are not listed here):

Leadership Council

- Will be responsible for setting and maintaining the strategic direction of EarthCube, and will identify what actions are within and outside the scope of that direction.
- Will assess and make changes to the governance structure as needed. (Will function as a mechanism to improve the structure.)

Confirmed Standing Committees and Teams

- Science Standing Committee
- Technology and Architecture Standing Committee
- Council of Data Facilities Standing Committee
- Science Funded Team / Technology Funded Team[2]
- Engagement Team
- Liaison Team

Working Groups

- Can be formed by committees, the Leadership Council, or emerge from the community.
- May be formed under the auspices of more than one committee, and thus can act as a mechanism for standing committees to work with each other.
- Can act as a way for the Leadership Council to delegate work to a subset of people with deep expertise and investment in a particular issue.
- Require formal approval; must have specific goals, timelines, and deliverables; may be allocated funds and staff support by the Leadership Council

Liaison Team

- The office will flag work in committees and connect with the Liaison Team to make sure efforts are connected with, and not duplicated by relevant external partners.
- Should have a steady implementation strategy to maintain proactive outreach to outside organizations, initiatives, and work outside of EarthCube.
- Will be heavily supported by the Project Office.

Engagement Team

- Serves a foundational function related to how core stakeholders are reached and involved in the organization.
- Responsible for maintaining proactive engagement of the EarthCube community.
- Will be heavily supported by the Project Office.

---

[2] These are written as “Funded Teams” in the charter and were referred to as such throughout the All Hands meeting. To clarify meaning, we have since renamed them as “Science Funded Projects Team” and “Technology Funded Projects Team.”

Several other issues were also discussed during the closing plenary session, and the following decisions were made:

Communication across Governance Organizational Units

It was decided that during the demonstration year, the Leadership Council should meet on a quarterly basis virtually. Standing Committees and Teams should have monthly teleconference meetings for those in leadership roles to foster communication across elements.

Science Committee/ Technology and Architecture Committee Integration

In order to promote communication and collaboration across these committees, there should be some cross-membership. Additionally, Working Groups with cross-membership will be encouraged.

Communication with NSF

To ensure unified messaging and priorities, it was decided that the Leadership Council should act as the single point of communication when advising NSF on behalf of EarthCube governance. It is understood that PIs on individual EarthCube awards will need to communicate directly with NSF program officers on matters related to their individual projects; this decision is not intended to subvert that need.

Engagement

It was decided that the Engagement Team should develop a structured communications plan within the first six months following the All-Hands Meeting.

Membership

For the first six months of the demonstration governance phase (and potentially beyond), it was decided that membership shall be free and open at an individual level in order to be as inclusive as possible as the community is developed. Institutional-level membership will be open initially as well. The structure of institutional membership will be determined in conjunction with institutions wishing to partner in order to best meet the needs of said institutions and EarthCube, and to ensure alignment with the Liaison Team. All membership models and processes may be subject to change pending review by the Leadership Council and Committee and Team leadership at a mid-year “Check & Adjust” meeting to be held in January, 2015.

Compensation for Leadership Council

No final decision was reached in regard to compensation for members of the Leadership Council and the Council Chairs. It was determined that an ultimate conclusion on this issue should be made by NSF within the next six months. While compensation remains an open issue, it was decided that rules for fiduciary responsibility, liability, etc., should be established by the Leadership Council.

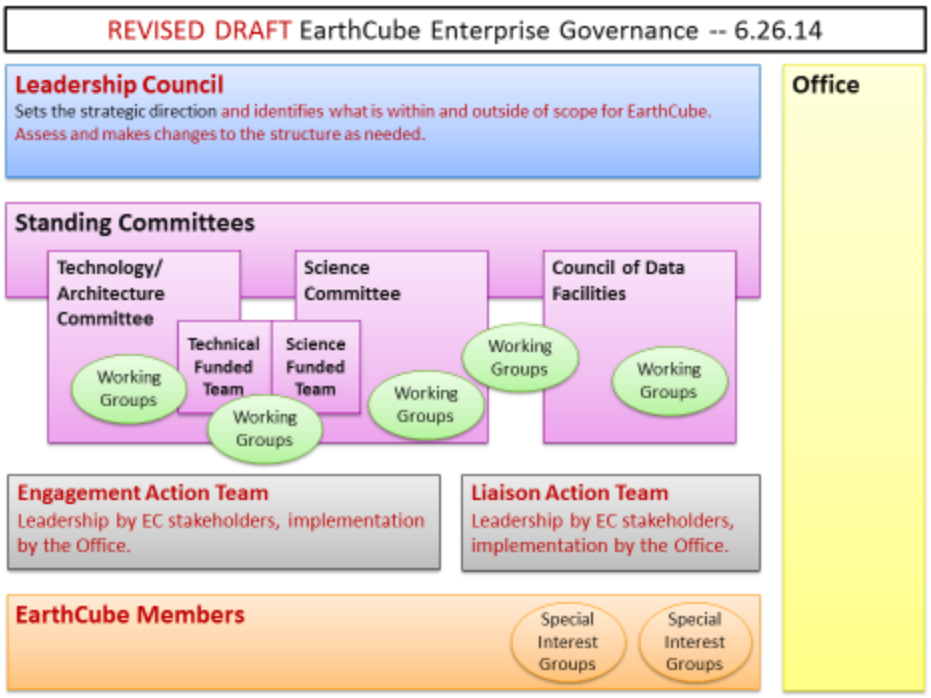


FIGURE 1. REVISED GOVERNANCE STRUCTURE, PRESENTED JUNE 26, 2014

Priority Issues for 2014–2015

In addition to a well-defined governance structure, one of the major intended outcomes of the All-Hands Meeting was to identify a set of priorities and actions to be carried out via the implementation of that structure. In preparation for the meeting, a list of priorities was identified using information gathered at the EarthCube Test Enterprise Governance Stakeholder Assembly workshops, as well as the more than two dozen NSF-funded End-User Domain workshops. In addition to this initial list of priorities, meeting attendees participated in a facilitated session in which additional priorities were developed. These lists were then combined, and participants were asked to vote for what they believed to be the most pressing issues that EarthCube governance should address during the demonstration year. During the closing plenary session, the selected priorities were presented to attendees with suggestions as to which committee or team should address them during the demonstration year. After some discussion, it was determined that during the first six months of the demonstration governance phase, initial responsibility of addressing these priorities will be divided amongst the committees and teams as follows:

- Refine the Scope and Vision for EarthCube (*Leadership Council*)
- Engage the scientific community and determine the structure of Working Groups (*Science Standing Committee)*
- Create a policy and/or strategy on EarthCube Data Management Plans (*Science Committee*)
- Consensus on EarthCube Architecture (*Technology and Architecture Standing Committee* in close coordination with the *Council of Data Facilities* and *Science Standing Committee*)
- Coordinate a testbed for EarthCube components (*Technology and Architecture Standing Committee*)
- Determine potential shared services through geosciences data providers (*Council of Data Facilities*)

- Identify and structure the Liaison program, focusing on collaboration with related activities in other agencies, industry, international, and other science domains (*Liaison Team*)
- Identify and share science success stories (*Engagement Team* in conjunction with the *Science Standing Committee*)

Governance Sessions and Processes

Governance sessions at the All-Hands Meeting were designed to gather participant feedback on the Charter, allow for real-time participant-driven changes, and facilitate group decision making on Charter modifications and other emergent issues. The first governance session, presented in plenary, focused on introducing the Charter and testing how the structure described would react to critical issues. This session, *Tackling the Tough Issues Facing EarthCube in 2014–2015*, was an interactive exercise designed to simulate the governance processes and dynamics to test the robustness of the proposed structure.

Participants in the simulation were asked to self-select into one of six roles, based on the committee structure detailed in the Draft Charter that was made public on May 6, 2014. Each participant was presented with one of two issues—one related to metadata, fields/disciplines, and technology, and the other related to credit and quality—and asked to consider the issue from the perspective of their role, first in separate meetings of committees and groups and then as combined groups with representation from each role. Participants were then put into groups with of each of the six roles represented, and asked to deliberate on the issues at hand using either a centralized or decentralized decision-making model. This exercise was followed by a facilitated discussion in which participants provided feedback on the governance structure.

This session resulted in many insightful observations and concrete suggestions for the proposed governance structure. The feedback can be grouped into several themes, summarized here:

Scope

Many comments from the session regarded scope. While there was strong interest among participants in clearly defining the scope of EarthCube and EarthCube governance, it was also recognized that some aspects of scope may need to remain deliberately broad to allow EarthCube to progress in unforeseen directions. A good deal of productive debate focused on striking the balance between a need for boundaries to drive effective decision making and community convergence on major issues, and the need to remain open to future opportunities for EarthCube. Additionally, participants expressed an interest in further clarifying the role of the NSF in EarthCube governance (e.g., whether participation in EarthCube governance is limited to NSF constituents, and what role NSF administration will play in governance processes).

Roles and Responsibilities

Participants also discussed the need for defining the roles and interactions of the governance components. At this early stage of the Demonstration Governance Framework's evolution, It has yet to be fully determined which groups will be empowered to make different types of decisions, and what procedures will be put in place to determine how decisions are ultimately made. Inter-committee relationships and responsibilities were another area of focus (e.g., what channels of communication exist between committees outside of representation on the Leadership Council?).

Decision Making

Discussions regarding decision making somewhat echoed the previous two issues, and focused on clarifying the decision-making process as written in the Demonstration Governance Charter brought forward at the meeting. Participants agreed that more work was needed in determining how final decisions would be made and by whom; for example, it was noted that the authority of the Leadership Council in the decision-making process required further definition. Other points of discussion addressed the role of budget in the organization, in particular the need for clearly defined financial protocols and workflows early on, as these could be influencing factors in decision making in the future.

Miscellaneous

Several important observations came out of the governance sessions that were not related to the themes above. For example, some participants noted that the structure of the committees was somewhat siloed in nature and not conducive to communication and collaboration among the committees and teams. Also, participants discussed the importance of balancing the considerable diversity in the backgrounds and fields that will be represented in the committees. If this is not done, there is a possibility that some groups may dominate while others may feel left out of the conversation. Finally, there was discussion about incentives for participating in EarthCube governance: It may not be clear to scientists and other participants what they will gain from their efforts, and participants agreed that this is a point that will have to be made explicit moving forward.

The opening governance session proved to be a valuable test of EarthCube as a system, and provided critical participant feedback on potential high-level issues that may arise. Subsequent sessions built on the outcomes of the morning's proceedings, and were designed to allow participants to provide further insight and to actively refine the governance model through facilitated discussions. In a plenary at the onset of Day 2 (*First Steps on EarthCube's Priorities for 2014–2015*), participants developed and refined a core set of priority issues that will become the charge of governance to address in the coming year. This was followed by an afternoon breakout session of about 30 participants (*Revising the Governance Model: Preparing for Next Steps*), in which groups of participants were assigned specific elements of the Charter and presented recommendations to the Test Governance project team. On Day 3, synthesized results from the previous sessions were presented and voted on in a final plenary *(EarthCube in Action: Next Steps for Tackling the Tough Issues).*

Architecture Forum Outcomes and Recommendations

A two-day Architecture Forum, organized by members of the Test Enterprise Governance IT Team, convened at the All-Hands Meeting with the primary goal of accelerating community convergence on an EarthCube architecture.

The group concluded that concrete decisions about the architecture of EarthCube cannot be made until its governance organization is in place; however, work is underway to seek consensus on the nature of the required architecture specification and how it will be used. Several points of general agreement emerged from the Architecture Forum plenary and breakout sessions, resulting in a set of recommendations to inform activities of the Technology and Architecture Standing Committee, tasked with the stewardship and development of an EarthCube architecture.

Organizational Recommendations

- The EarthCube architecture framework will be a constant work in progress, managed by an evolutionary process to adapt to changing technology and community requirements.
- The scope of the architecture must be determined through broad engagement of the Earth Science research and education community to meet their needs.
- The architecture framework can be used to identify gaps in capabilities and to establish development priorities.
- The business model for funding and maintaining production infrastructure should be treated as distinct from that supporting technology research and cutting edge development.
- In recognition of the need for EarthCube Architecture to be highly flexible in the long term, the term “Reference Architecture” should be avoided in committee communications, in favor of “EarthCube Architecture” or “Evolving EarthCube Architecture.”

Design Recommendations

- EarthCube architecture must support and facilitate continuous change to adapt to new research techniques and priorities and technology innovation.
- EarthCube architecture should provide a guide to assess what technology is or could be part of the system.
- A 'top down' design framework is useful to guide system development and management to assure that component Building Blocks work together, but the system must be responsive to 'bottom up' input based on actual practice that emerges in the community.
- EarthCube components should be modular and loosely coupled, reducing the barriers to plugging in new components.
- EarthCube will be a system of systems, and therefore must incorporate competing approaches, existing operational systems both within and outside of the NSF portfolio.
- As an NSF system of systems operating in the context of a larger Earth Science Research infrastructure, the architecture will need to focus on the gateways (interfaces and information exchange agreements) that connect components and enable interoperability.
- EarthCube architecture must occupy middle ground between the uncoordinated technology-Bazaar model and the single monolithic system model.

Architecture Viewpoints

Information system architecture requires specification from a variety of viewpoints, each providing different interrelated perspectives of an overall system. Part of the architect’s role is to make sure any set of viewpoints is consistent and not contradictory, though there is not, however, a universal standard of which viewpoints are to be used.

Based on presentations and outcomes from the EarthCube Architecture Forum, the Test Enterprise Governance IT Team was able to identify a preliminary set of useful viewpoints that can be considered in developing EarthCube architecture, though additional viewpoints will likely be incorporated in future discussions. In this case, the team focused on a set of four viewpoints defined in the ISO/IEC standard Reference Model for Open Distributed Processing (ISO/IEC 10746, see http://www.rm-odp.net/).

- The ***Enterprise Viewpoint*** includes the scope and objectives of EarthCube, the users and stakeholders, example Use Case scenarios, and top-level enterprise components (e.g., scientists, cyberinfrastructure, and data facilities).
- The ***Information Viewpoint*** addresses the semantics of the information and data which is relevant to EarthCube. Key information models (e.g., geographic), are defined with reference to international standards as available.
- The ***Services Viewpoint*** includes interfaces and workflows pertinent to EarthCube as a cyberinfrastructure based on a service-oriented architecture.
- The ***Engineering Viewpoint*** identifies key types of components required to support the deployment, management, and integration of the services and data.

Revised Critical Functions of the Technology and Architecture Standing Committee

1. Ensure the explicit connection between the scientific process and technical function, including coordinating testbeds and other mechanisms for development of cyberinfrastructure components, for use in science use cases.
2. Seek alignment of EarthCube funded projects to foster integrated technology to meet end-user requirements and create new capabilities to enable transformative geosciences research.
3. Provide stewardship of an EarthCube architecture.
4. Address how EarthCube will accommodate or incorporate different standards and standards bodies.
5. Identify gaps in coverage of needed cyberinfrastructure capabilities, and determine recommendations on how to fill them.
6. Monitor technical requirements with the goal to ensure that EarthCube is meeting end-user needs.
7. Serve as an emissary between software developers, the science community, and infrastructure providers, as well as educators.
8. Encourage, engage, and enable the next generation of EarthCube technology leadership.
9. Coordinate with the Technical Funded Team.

Architecture Forum Sessions and Process

With a demonstration EarthCube governance structure now taking shape, the EarthCube community was ready to take its first steps toward convergence on an architecture, one of the key deliverables for Year 2.

The Architecture Forum convened with the primary goal of accelerating community convergence on an EarthCube architecture. To begin that process, the community would need to reach initial agreement on three basic elements: the terminology to be used in discussion, the purpose of an EarthCube architecture, and a broad idea of what an EarthCube architecture might look like with regard to scope and granularity.

A secondary goal of the workshop was to take first steps in laying the groundwork for year 2 activities of the Technology and Architecture Standing Committee, to be put into place October 1, 2014 as part of the emerging Test Enterprise demonstration governance structure.

The forum resulted in several key accomplishments:

- Made significant progress toward a shared common understanding of what constitutes an ongoing architecture.
- Identified critical issues of concern to community members regarding an EarthCube architecture.
- Assembled representatives from all of the funded Building Blocks to determine potential alignments, interfaces, and architecture needs, establishing an inventory of emerging architecture based on the EarthCube funded projects. (See Figure 3)
- Collaboratively revised elements of the Technology and Architecture Committee Functions as outlined in the EarthCube Test Enterprise Governance Charter.
- Through the use of a collaborative "living glossary,"[3] created a basis for future communication across the EarthCube technology community.

Introductory Presentations

The initial 90-minute plenary session had three main components, each supporting the session objective of establishing a framework for Building Blocks projects to fit into the Conceptual Designs. After introductory remarks, the National Science Foundation presented its own vision of EarthCube architecture with regard to scope, challenges, and strategy. In terms of scope, architecture should be determined by the community through broad engagement and must meet the needs of the Earth Science research and education community. Major challenges to these efforts include competing ideas within the EarthCube community about the best approach to developing an architecture; the issue of having to integrate legacy systems into an architecture; and the diversity and complexity of information in the Earth sciences. Finally, the NSF's strategies for moving forward involve architecture development based on end-user needs, an incremental process based on merit, feasibility, and competence, and three defined phases: conceptual design, refinement, and implementation.

A panel discussion immediately followed, in which representatives from OceanLink, iPlant, and GEOSS discussed use cases, best practices, and lessons learned from their own experiences with architecture development. Several key learnings emerged from this discussion. Panelists envisioned community infrastructure as a system of systems that is continuously undergoing evolutionary development and must adapt to and satisfy user needs in a community-driven way. Because constituent systems will have their own architecture, EarthCube must integrate these by focusing on the interfaces that link them. To do this, panelists noted the importance of employing both bottom-up and top-down approaches in developing an ongoing architecture.

Principal Investigators from the two current EarthCube Conceptual Design awards (*Developing a Data-Oriented Human-Centric Enterprise Architecture for EarthCube*, and *Enterprise Architecture for Transformative Research and Collaboration Across the Geosciences)* finished the session with a preliminary progress report, outlining their current process and draft design thus far, with a focus on what they viewed as requirements and how they envisioned the design could eventually be used. These

[3] Glossary available at: http://bit.ly/1q8NWqe

reports served to provide attendees with a vision of emerging technology in EarthCube, in an effort to frame discussion for the following day's activities.

A Common Understanding of Architecture

Day 2 consisted of two interactive 90-minute sessions with an aim to move toward a shared understanding among participants of what is meant by a "reference architecture" in relation to EarthCube, and to plan for an architecture that will support EarthCube science objectives (as determined via the 2012–2014 End-User Workshops).

Through the use of a collaborative notes system, a dynamic glossary, and open discussion, attendees reviewed the previous day's activities and came to agreement on a common description of an EarthCube architecture: A high-level view of a system that can used as a guide to assess whether some technology fits or doesn't fit. The architecture as described would be akin to a reference book for a discipline, used to identify gaps and establish priorities.

Coordinating the Building Blocks

In an effort to create an inventory of the Building Blocks and identify potential collaborations (Fig. 1), representatives from each award team presented a 5-minute lightning talk to outline their overall progress and emerging components thus far. These presentations provided an inventory of the EarthCube component portfolio that is currently under development.

Component Inventory: Potential Interactions between EarthCube Building Blocks and RCNs

	GeoWS	ODSIP	GEOSOFT	BCube	DisConBB	OceanLink	DeepDive	CINERGI	ES Bridge	RCNs
GeoWS				X				X		
ODSIP	X			X				X		
GEOSOFT				X				X	X	
BCube	X							X		
DisConBB	X	X								
OceanLink								X		
DeepDive										X
CINERGI	X	X	X	X		X				X
ES Bridge	X									

FIGURE 2. INVENTORY OF EMERGING ARCHITECTURE BASED ON THE EARTHCUBE FUNDED PROJECTS

A subsequent breakout exercise called on participants to form small groups to develop a scenario for assembling components from this portfolio, along with other existing components, to perform some specific function that would be useful to them. Groups were asked to produce a block diagram showing how the components would work together. These diagrams were subsequently drafted and summarized, and can be viewed in Appendix 1: *Use Case Scenarios*). This series of use cases identified multiple component types, the commonalities of those components, and potential interfaces between them, including:

- Model Validation with Data (BCube, CINERGI, DisconBB, Earth System Bridge, GeoWS, ODSIP)
- Integration of Ocean Floor Data (CINERGI, DisconBB, GeoWS, OceanLink, ODSIP)
- Broker Architecture (BCube, Earth System Bridge, GeoWS, ODSIP)
- Field Data Acquisition (CINERGI, GeoSoft, GeoWS, OceanLink, ODSIP)
- Mining Resource Inventory (CINERGI, GeoDeepDive)

Through this exercise, workshop participants were able to get a high-level view of current Earthcube Portfolio technology efforts and begin identifying gaps that will eventually need to be filled. The use cases presented in this forum will be revisited as members of these projects begin collaborating via the Demonstration Governance Structure, particularly in the Technology and Architecture Standing Committee and the Technical Funded Team. The exercise was instructive for exploring the role that an evolving architecture could play to assist the design process, both to assemble components for science workflows, and for system design and development priorities.

Representatives from the Building Block teams create diagrams detailing potential cross-project collaborations.

Developing the Technology and Architecture Committee

The final session was dedicated to the establishment of a Technology and Architecture Standing Committee as part of the EarthCube Demonstration Governance framework. This committee, which will go into place October 1, 2014, will be tasked with coordinating development of the technology that is part of EarthCube, including stewardship of the ongoing architecture.

Through facilitated discussion, participants reviewed the list of Critical Functions for the Technology and Architecture Standing Committee as laid out in the EarthCube Demonstration Governance Framework. A short breakout exercise asked participants to create a list of five work items and responsibilities that should be represented within this framework; upon regrouping, participants and facilitators made revisions to the Critical Functions portion of the Charter document based on general consensus of the group.

Engagement & Community Building Outcomes and Recommendations

Outcomes of the All-Hands Meeting included several recommendations for next steps related to engagement and community building, in order to successfully move into and carry out the Demonstration Phase. The first of these was to inform the larger EarthCube community of the outcomes of the All-Hands Meeting and how they can get involved in the project moving forward. In evaluating feedback from previous meetings, and outreach strategies, it was found that there is a perception in the community that only those individuals who can attend the meetings/workshops/etc. can take an active role in EarthCube.

In order to respond to this feedback, it will be necessary to ensure there are adequate ways in which the community can get involved in the governance process, especially during the initial elections of Leadership Council members and initial organization and leadership selection of the Standing Committees and Teams prior to the planned Demonstration Governance start date of October 1, 2014. As a first step in this, an EarthCube Newsletter was sent out shortly after the All-Hands Meeting, outlining the outcomes of the All-Hands Meeting and providing the following preliminary timeline for

virtual involvement. (To ensure balanced representation from all geosciences domains, however, the deadline for selection of at-large members has since been extended.)

- July 17, 21 and Aug. 5: Webinars on Standing Committees and Teams
- Late August: First Standing Committee and Team virtual meetings
- August-September: Nomination and selection of At-Large Leadership Council members
- October 1: Deadline for selection of Committee and Team representatives to Leadership Council

The Newsletter was followed up with social media posts, Listserv messages, email outreach to past Workshop participants, and other venues outlining the same information. In addition to this information, the community was given the opportunity to fill out an expression of interest form to join a Standing Committee or Team. An open question and answer forum was provided at the end of each webinar to allow the community to have any queries and concerns addressed[4].

Another outcome of the All-Hands Meeting was to the need to launch a renewed community outreach program aimed toward sufficiently populating the Demonstration Governance structures (*Science Standing Committee, Technology and Architecture Standing Committee, Engagement Team, and Liaison Team)* in order to effectively test the proposed Charter and allow for elements to be adjusted to better meet community needs. There was an extra effort to solicit the participation of academic geosciences researchers for the Science Standing Committee for three reasons: 1) the NSF-funded academic geoscientist community is the primary stakeholder for EarthCube, 2) the Science Standing Committee holds the important task of determining the scientific direction of EarthCube in the Demonstration Phase, and 3) there was a lesser presence of academic geoscientists at the All-Hands Meeting in comparison to the CI community.

The revamped community engagement plan focused on leveraging already-engaged EarthCube community members to distribute a call for participation to their distinct communities, as well as provide our team with specific names of individuals within their community who they feel would be a positive addition to the demonstration governance. Community members at institutions/organizations such as UCAR, Scripps, Oregon State, WHOI, and AMS connected our team with their respective communications team/director and the involvement information outlined above was relayed to their communities.

Finally, the recommendation for greater engagement and collaboration with and among the EarthCube funded projects was a major outcome of the All-Hands Meeting. During the meeting, an ad hoc lunch session was organized for all funded EarthCube projects, with participation from the Governance team and NSF, which focused on the role of the projects within the Demonstration Governance as a means to foster communication across and with the Standing Committees. A major outcome of the lunch meeting was recognition of the need for a plan for regular communication, coordination, and collaboration across current and future EarthCube funded projects. Concrete ideas that might possibly contribute to this plan include:

- Hack-a-thon/'Radical Co-Location' to bring groups together in person to identify and carry out work that needs to be done

[4] Recording of Community Webinar at:
https://www.youtube.com/watch?v=fyQ5RgM8pjo&list=UUAZaYvuagNoTRK7GPMUnUQQ

    - Technical, active, high-bandwidth dialogue
    - Gap analysis to inform future NSF solicitations
    - Making connections back to End-User Workshop outcomes
- A Metadata Alignment workshop
- An internal review process to analyze the entire system design to evaluate how all the technical pieces fit together

Meeting Metrics

Attendance

The All-Hands Meeting brought together nearly 150 members of the geo- and cyber-science communities, representing over 90 academic institutions, government agencies, non- and for-profit organizations, and consortia. Attendees included both old and new faces to the EarthCube community—people who had never attended an EarthCube meeting before as well as people who had been at the first EarthCube Charrette in 2011. Each EarthCube Building Block, Conceptual Design, and Research Coordination Network project had at least one representative in attendance. Furthermore, representatives from each Assembly Stakeholder workshop were present including the Data Facilities, Information Technology, Computer Science, Free and Open Source Software, Professional Society, and End-User Geoscientist communities. A complete list of projects, institutions, and organizations represented at the All-Hands Meeting can be found in Appendix 2.

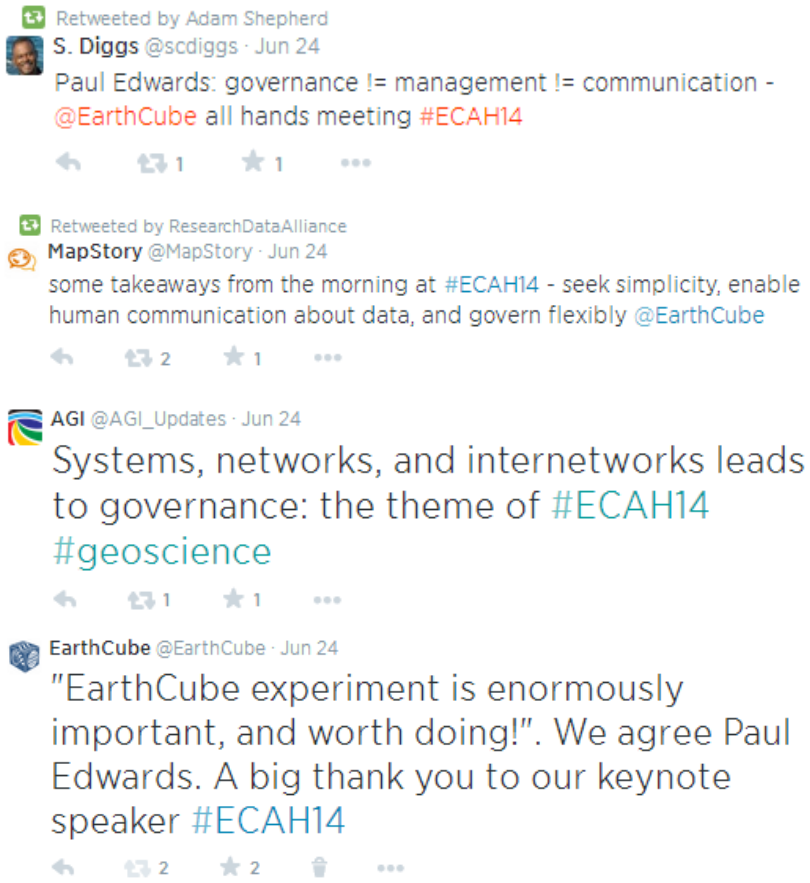
Sessions

In order to reach meeting goals, the agenda included structured plenary and breakout sessions or business meetings as well as on-the-fly sessions, group meetings, and ad hoc discussions/presentations. The EarthCube Test Enterprise Governance Team and Meeting Organizers received over 25 proposals and were able to accommodate all within the agenda. (See Appendix 1 for a complete list of community proposed sessions.) Sessions ranged from demonstrations of technical components (both EarthCube and non-EarthCube), discussion panels on addressing software and cyberinfrastructure issues, and discussions about the social side of cyberinfrastructure efforts such as governance and workforce development. In addition to the sessions, two special events took place. First, an “Intro to EarthCube” presentation to kick off the meeting was held in order to detail the history, goals, and current developments of EarthCube for those who were new to the project or needed a refresher. Second, an evening Poster Session highlighted the work of over 30 authors/teams in the EarthCube community.

Social Media

The EarthCube Test Enterprise Governance Project Team assembled a Twitter Team to develop and carry out a social media plan with the primary goal of engaging and informing meeting attendees (and the larger EarthCube Twitter community) of all that was happening at the All-Hands Meeting throughout the day. The Twitter Team created the hashtag for the meeting (#ECAH14), using it in all related tweets. Furthermore, attendees were encouraged use #ECAH14 as well in order to capture all of the sessions, presentations, conversations, and questions that arose throughout the meeting in an open, virtual platform. This live stream of thoughts, reactions, and photos allowed attendees to share their ideas and learn what others were doing and thinking throughout the event, and fostered a virtual dialogue between disparate activities and discussions.

The Twitter Team's strategy was a resounding success. The #ECAH14 hashtag generated 400+ tweets and retweets from over 30 different users (both individuals and organizations/projects). Highlights of the All-Hands Meeting Twitter feed included thoughts on Dr. Paul Edwards' presentation about the history of infrastructure and how it applies to earth sciences, insights/questions generated by technical sessions, and positive reactions to the fact that EarthCube is action and how it i moving forward.



Excerpt of a Twitter conversation regarding Paul Edward's presentation

Next Steps: Moving Forward into Year 2

The Year 2 demonstration phase will test the effectiveness of the proposed framework and allow for elements to be changed to better meet community needs. It will begin by populating committees and teams, and finalizing leadership and decision-making processes to move forward on community-selected priorities, including identifying science drivers, coordinating emerging technical elements, and coming to convergence on system architecture. A January mid-year review will assemble these groups to analyze the effectiveness of the framework thus far and make adjustments as necessary. The general timeline and next steps for this process is as follows:

- Late August: Initial Committee and Team Virtual Meetings
- September to December: Working Committee and Team Virtual Meetings
- October: Demonstration organization in place

- January: “Check & Adjust” mid-year review—Testing of Governance process with structure changes and refinements as needed
- February to June: Initial development of an EarthCube Architecture Framework
- June: Review of Governance process with recommendations for adjustments as needed
- June: Results presented at the 2015 EarthCube All Hands Meeting

APPENDIX 1: COMMUNITY PROPOSED SESSIONS

THE VHUB COMMUNITY CYBERINFRASTRUCTURE—SHARING AND GROWING

Convener: Patra, Abani (SUNY at Buffalo)
Co-Conveners: C. Connor, L. Connor, G. A. Valentine, M. D. Jones, S. Gallo, P. Webley
Session type: Workshop/Hack-a-thon
Description: A fundamental goal of cyberinfrastructure in the geosciences is to enable integration of multidisciplinary and computational thinking into research and applications. Ideally this cyberinfrastructure provides a mechanism for globally collaborative research and development of computational models of geological processes and their integration with complex geospatial, observational, and experimental data. In order to achieve these aspirations, it is critical to develop resources, within the cyberinfrastructure itself, to promote the training of the next generation of geoscientists in management of data, modeling and simulation.

The session will start with a broad introduction of the successful VHub (Vhub.org) collaboratory that has a significant fraction of the volcanology community using it for various services including but not limited to model and data sharing, online simulation, project based collaboration and community organizational efforts. We will follow with more participatory exercises on using the VHub platform for shared projects including modeling software development.

In this session we will focus on the use of cyberinfrastructure platforms, like VHub, to enhance the academic experience of students, convey concepts in code verification and validation, and to train new researchers in modeling and simulation. We are interested in highlighting innovative methods to ease students and other learners through the process of development of new research skills: from running code and manipulating outputs, considering model verification and validation, through to modifying code and contributing new code to the research community. Development of this perspective is essential for EarthCube, not only for students but to increase the flexibility and long-term sustainability of cyberinfrastructure and software used in the geosciences, as in other disciplines.

GEOWS, CINERGI, BCUBE BUSINESS MEETING

Conveners: Ahern, Tim (IRIS)
Session Type: Business Meeting

DISCONBB MEETS IOOS MARINE IN-SITU DATA

Convener: Arctur, David
Co-Conveners: Emilio Mayorga (University of Washington), Ethan Davis (Unidata), Alva Couch (Tufts University)
Session type: Business Meeting
Description: The Building Block on Integrating Discrete and Continuous Data (DisConBB) is developing a common information model and tools for converting between gridded and time series data, initially for the atmospheric and hydrologic domains. This session is for discussion to develop and agree on this common information model and tools for working with WaterML 2 and netCDF/CF, using OGC SOS, THREDDS/OPeNDAP and related web services. This supports DisConBB's end goal to have something that can be adapted for other domains, such as oceans, solid earth, and cryosphere. Will include

discussion on existing, relevant efforts that can inform DisConBB's solution, including the Unidata Common Data Model Discrete Sampling Geometry (CDM DSG) feature types and the IOOS SOS / SWECommon profile and associated server software implementations that bridge the netCDF/CF/CDM DSG/THREDDS and OGC SOS/O&M/SWECommon/SensorML approaches.
A key issue is that while there are a plethora of options for "mapping data" between formats, the tool chains we have available for analyzing data in the new formats don't necessarily interoperate well with the mapped data. We are interested in how these mappings would be used, in what tools, and for what outcome. One large issue is how to generate NetCDF files of WaterML data that can then be applied in compelling use cases. We can get the NetCDF, but then tools fall short of being able to do useful things with it.

ADDRESSING DATA HETEROGENEITY IN EARTHCUBE INFORMATION: A SEMANTIC BUILDING BLOCKS & CYBERINFRASTRUCTURE PERSPECTIVE

Convener: Berg-Cross, Gary (SOCoP)
Co-Conveners: Pascal Hitzler (Wright State University), Kerstin Lehnert (Columbia University), Peter Wiebe (WHOI)
Session Type: Panel Discussion
Description: The EarthCube (EC) community is currently engaged in a range of activities including creation of building blocks to handle Big Data issues and meet critical needs to find, organize, interpret, and share appropriate data & resources for specific research activities across EC disciplines. Current sustainable cyberinfrastructure employs a wide assortment of traditional, but useful standards, methods, and tools. It also is starting to include aspects of the still maturing and promising set of tools, methods, and standards associated with Semantic Technologies. These semantic initiatives, which are being incorporated into traditional infrastructure along with supporting practices, are aimed at dealing more effectively with stubborn, central challenges that arise from the inherent complexity and heterogeneity of data that is structured and unstructured, and based on independently developed domain models. Such direct, but tractable approaches developed to address semantic interoperability, are aimed at enabling better data integration, analysis, and dissemination to support cross-disciplinary research.
This panel will provide an opportunity to review & share current progress and experience with semantic technology approaches, including linked data, the use of lightweight, opportunistic methods, bottom-up & top-down approaches, and related infrastructure. Taken together, how important are semantic building blocks and semantic technology drivers for addressing Big Data variety challenges and the handling of diverse data and infrastructure issues? What are measures of success and risk? The panel discussion is expected to cover new opportunities and issues such as: schema mapping, approaches to semantic annotation and terminology, selection of appropriate knowledge representations, development and use of ontologies and modular ontology patterns, and integration of semantic technology and its methods with traditional technology. A question for the panel: what extant work and approaches can be leveraged and how might this be done as part of plans going forward with the EarthCube community, its Research Coordination Networks, and building blocks efforts?

PROTOTYPE ACTIVITIES WORK GROUP BUSINESS MEETING

Convener: Diggs, Steve (Scripps)
Session Type: Business Meeting

WORKFORCE DEVELOPMENT PANEL
Convener: Downs, Robert (CIESIN, Columbia University)
Co-Conveners: Robert S. Chen (CIESIN, Columbia University), Leslie Hsu (LDEO, Columbia University), Kerstin Lehnert (LDEO, Columbia University), Erin Robinson (Foundation for Earth Science), Ilya Zaslavsky (University of California, San Diego).
Session Type: Panel Discussion
Description: The EarthCube community is faced with various challenges, opportunities, and choices for enabling Earth science. Technology is evolving rapidly and new tools and techniques are being created and enhanced to enable community members to do more individually and collaboratively. Workforce development opportunities can assist community members in their efforts to improve their capabilities and knowledge. A panel will describe workforce development opportunities that are available for the Earth science community and engage in a discussion with the audience on how the EarthCube community can improve by leveraging these opportunities.

CRAWLING THE WEB FOR EARTHCUBE
Convener: Duerr, Ruth (NSIDC)
Co-Conveners: WenWen Li (Arizona State University), Siri Jodha Khalsa (University of Colorado)
Session Type: Workshop/Hack-a-thon
Description: Structured advertisements for web services and data exist all over the web often in forms such as OpenSearch descriptions, OGC web service capabilities documents, WSDL documents, OAI-PMH metadata feeds and other metadata files, and semi-structured data in semantic web languages. The data and services pointed to by these documents are central to scientific analysis and decision making. Various tools now exist to allow researchers to simply fill out web forms to advertise (i.e., “publish”) the existence of their data in one of these discoverable forms. Technologies such as large-scale web crawling can be used to actively search, discover and aggregate these “ads” and provide that information to catalogs and brokering services. Several EarthCube building blocks are funded to use these technologies to improve the discoverability of distributed web resources. In this 90-minute workshop these plans will be discussed, interfaces fleshed out, and feedback on the types of data and services of most important to the community solicited.

SOFTWARE STEWARDSHIP TEAM BUSINESS MEETING
Convener: Gil, Yolanda
Session Type: Business Meeting

SOFTWARE SUSTAINABILITY IN GEOSCIENCES
Convener: Gil, Yolanda (University of Southern California)
Co-Conveners: Christopher J. Duffy (Department of Civil and Environmental Engineering, Penn State University), Chris Mattmann (Department of Computer Science, University of Southern California and NASA/JPL), Scott D. Peckham (Department of Hydrologic Sciences, University of Colorado), and Erin Robinson (Foundation for Earth Science).
Session Type: Panel Discussion

Description: This panel will discuss new findings regarding the requirements of geoscientists to manage their software, best practices for software sharing and reuse, and important social considerations surrounding scientific software.

TEN SIMPLE RULES FOR THE CARE AND FEEDING OF SCIENTIFIC DATA

Convener: Gil, Yolanda (University of Southern California)
Session Type: Technology Presentation
Description: This presentation will give practical guidelines for scientists to take care of their data. The guidelines are captured as ten simple rules, and are based on a recently published paper (10.1371/journal.pcbi.1003542) that resulted from a collaborative workshop on data citation and reuse.

ALIGNING REPRESENTATIONS: PRACTICAL EXPLORATIONS OF METADATA, ONTOLOGIES, AND SEMANTIC WEB FOR EARTHCUBE

Convener: Hitzler, Pascal (Wright State University)
Co-Conveners: Joel Cutcher-Gershenfeld (University of Illinois at Urbana Champaign), Yolanda Gil (University of Southern California), Krzysztof Janowicz (University of California, Santa Barbara), Scott Peckham (University of Colorado, Boulder), Steve Richard (Arizona Geological Survey), and Ilya Zaslavsky (University of California San Diego).
Session Type: Workshop/Hack-a-thon
Description: Metadata, ontologies, and conceptual models for representing information play a role across many current EarthCube efforts. This shared representational theme thus has the potential to serve as a catalyst for technology integration across EarthCube, by starting to discuss possibilities for aligning these representations, or for mapping between them.

This workshop is proposed by representatives from several current EarthCube awards and the idea rose out of corresponding discussions at the EarthCube Awards Assembly in February 2014. The workshop shall map out the representational choices made by different technical approaches within EarthCube. Hands-on, workshop participants will identify practical technological solutions for obtaining alignment or mappings between representations at EarthCube scale, and will identify major technological or social obstacles which may prevent such alignment or mapping. Given enough time, the workshop participants will jointly attempt to sketch concrete technological solutions for metadata, ontologies, and conceptual models which would support the EarthCube vision.
The workshop will start with briefings on current representational choices made by EarthCube awards, focusing on design choices and their underlying rationales. Small breakout groups will then work on possible mappings or alignments between a subset of the presented representations. Results from these breakout groups will be reported to all participants, groups would be further combined into larger groups to incorporate further representational choices in the discussion. A plenary session will combine and record the insights made and resulting recommendations.

OCEANLINK BUSINESS MEETING

Convener: Hitzler, Pascal (Wright State University)
Session Type: Business Meeting

DATA ASSIMILATION FRAMEWORKS IN GEOSCIENCES
Convener: Kelbert, Anna (Oregon State University)
Session Type: Panel Discussion
Description: The ultimate goal of any modeling in geosciences is gaining a better understanding of the physical, chemical and biological processes going on in the Earth's interior and environment. This understanding is best attained through a comparison of the modelled results with measured data. A formalization of the procedure, in which measurements are driving the modeling to ensure compatibility, is known as data assimilation. Since most Earth processes are tightly interconnected, multi-physics data assimilation may sometimes be desirable to get the most of the data we have, integrating all measurable information into a merged 4D snapshot of the Earth.
I would like to use this session to get a better understanding of the following issues.

1. Which areas of geosciences would benefit from more streamlined, multi-physics data assimilation?
2. Is a general and easy to use framework for multi-physics data assimilation practical?
3. What is our community experience so far with data assimilation frameworks?
4. What are the scientific and technical challenges, and (how) can they be overcome?

Finally, I'd like to get a feel for the community consensus on the role of multidisciplinary data assimilation in EarthCube, as the program evolves.

THE BCUBE BROKERING FRAMEWORK
Convener: Khalsa, Siri Jodha (University of Colorado)
Co-Conveners: Stefano Nativi (CNR), Ruth Duerr (NSIDC), Steve Browdy (OMS Tech), Jay & Francoise Pearlman (J&F Enterprises), Oscar Schofield (Rutgers), Scott Doney (WHOI)
Session Type: Technology Presentation
Description: We will demonstrate the data and services provided by the BCube Brokering Framework in support of its science scenarios. We will show how the resources of the Brokering Framework, which is hosted in the Cloud, can be accessed via disciplinary portals or desktop tools, using standard web protocols or platform independent APIs.

BCUBE PROJECT DISCUSSION AND BUSINESS MEETING
Convener: Khalsa, SiriJodha (University of Colorado)
Session Type: Technology Presentation and Business Meeting

MIDDLEWARE GOVERNANCE
Convener: Khalsa, Siri Jodha (University of Colorado)
Co-Conveners: Stefano Nativi (CNR) and Jay Pearlman (J&F Enterprises)
Session Type: Panel Discussion
Description: Middleware can significantly simplify creation and operation of distributed systems, as well as providing a much more efficient means of integrating legacy systems with new technology. For example, brokering middleware provides mediation and transformation services to simplify data discovery, evaluation, access and use, working as a third-party tier in a three-tier architecture (extending the Client-Server paradigm). This introduces the clear need to govern and manage a middleware tier from many perspectives such as interface, stability and sustainability.

Effective middleware governance, as part of an overall governance approach, has the potential to support longer-term development under a variety of funding models, to simplify and standardize access models and assist in establishing a basis for development and operation of an EarthCube information system. It is not, however, clear what the best practices for this governance are and how those practices shift in response to different funding and ownership models, under different architectures or as standards change. To ensure sustainable, stable development, an effective model for the governance and reuse of such middleware must be agreed upon. This will be of value not only to interoperability architects and to developers (who can plan integrated systems assuming the continued use and support of middleware) but also to system managers and end users. We propose to consider this in a session on middleware governance at the EarthCube All-Hands meeting, in collaboration with the Test Governance team.

EARTHCUBE CLOUD COMMONS: GEOCLOUD WORKSHOP

Convener: Law, Emily (Jet Propulsion Laboratory)
Co-Convener: Wenming Ye (Microsoft Research)
Session Type: Workshop/Hack-a-thon
Description: The EarthCube Cloud Commons (ECC) working group's objective is to investigate issues related to adoption of cloud computing by the NSF, educate and facilitate adoption of cloud computing by the EarthCube community. A pilot project is currently underway in developing a prototype cloud Virtual Machine (VM) repository with an online cloud portal that provides guidance on how to create cloud instances using ECC. A demonstration of the prototype and tutorial on how to use cloud portal will be given at the workshop. The workshop also provides an opportunity to engage the EarthCube community at large, and provides a forum for ECC discussions and inputs.

INVENTORY OF EARTHCUBE RESOURCES: CINERGI ARCHITECTURE AND WORKFLOW

Convener: Malik, Tanu (University of Chicago)
Co-Conveners: Ilya Zaslavsky (University of California San Diego), Steve Richard (Arizona Geological Survey), David Valentine (University of California San Diego), Jeffery Grethe (University of California San Diego), and Burak Ozyurt (University of California San Diego).
Session Type: Technology Presentation
Description: The CINERGI project is constructing a community inventory and knowledge base on geosciences information resources to meet the challenge of finding resources across disciplines, assessing their fitness for use in specific research scenarios, and providing tools for integrating and re-using data from multiple domains. Constructing such a knowledge base can be challenging since data may need to be harvested from multiple sources, but more so data needs to be extracted and cleaned for easy consumption.

This technology presentation will describe in-depth the architecture, tools, and services used in creating the CINERGI repository. In particular, it will describe the how multiple information extraction pipelines are assembled using GeoPortal, a combination of customized operators/parsers, and MongoDB to curate harvested resource descriptions in a scalable way. Given the heterogeneity of the data it is often the necessary to do post-mortem analysis of the outputs. We will describe a comprehensive provenance framework that captures extraction details from these components and answers vital questions about a resource.

This technology presentation is targeted for cyber-infrastructure specialists who manage large information extraction pipelines, and curators and users who have queries relating to provenance of resources.

SIMPLIFYING SCIENTIFIC DATA MANAGEMENT THROUGH SAAS
Convener: Malik, Tanu (University of Chicago)
Co-Conveners: Kyle Chard (University of Chicago)
Session Type: Technology Presentation
Description: Scientific projects must deal with increasingly large and diverse data. These data encompass many files, database tables, and linked data; feature different data models and file formats; and are located in different storage systems and institutions. When data are simple and few, a researcher may track them in a notebook or spreadsheet; the data themselves may simply be stored in a common directory. For example, files related to a specific simulation run may be stored in a directory named by run number; files related to a specific field site may be stored in a directory named for that site.

The fact that current data management practices do not meet the needs of scientists is evident in that scientists spend significant amounts of their research time managing data rather than doing science, with self-reported values of 90% being common. With trends towards “big data” and the associated increases in data size, variety, and complexity, data management problems will grow yet worse. Simply keeping track of the data can quickly become an unmanageable task —greatly complicating research processes.

This technology presentation will demonstrate services for the entire data life cycle. It will first demonstrate Globus services of data movement, synchronization, and sharing. It will then focus on data publication and discovery capabilities that are being added to Globus service. These new features make it simple to describe, curate, and preserve data at desired levels of durability. They also enable rich discovery by making it possible to search, browse, and access large published data sets. Finally, we will demonstrate services that make it easy for users to track data, its transformations and view the entire provenance trail of the data. This technology presentation will be hands-on; demonstrating and walking through with users how scientific data management tasks can be simplified drastically throughout-sourced services.

COUNCIL OF DATA FACILITIES BUSINESS MEETING
Convener: Ramamurthy, Mohan (Unidata)
Co-Conveners: Kerstin Lehnert (Columbia University), Don Middleton (UCAR)
Session Type: Business Meeting

SOCIAL DIMENSIONS OF EARTHCUBE DATA AND INFRASTRUCTURE
Convener: Slota, Stephen (University of California, Irvine)
Co-Conveners: Siri Jodha Khalsa (University of Colorado), Geof Bowker, (University of California, Irvine)
Session Type: Panel Discussion
Description: The development of new data infrastructures for collaboration and potential large-scale analysis is as much a cultural and social exercise as it is technical. Organizational, cultural, and political

issues move to the forefront as various facets of wide-scale uptake, disciplinary division and data practice become the largest challenges to infrastructure adoption and use. When science or scientific outcomes are published, made open or accessible, or otherwise released into the world they bear the potential for substantial cultural, social and political effects. What consequences are there to publicly opening data and providing the basic tools for their interpretation and use? What are the social and policy outcomes of the publication of data, models and analysis? Organizational, technological and disciplinary cultures play a significant role in the use and effective re-use of scientific databases, models and tools. Social scientists have long studied the movement, storage, retrieval and re-usability of scientific data, but their role changes based on the projects and scientific communities engaged in various collaborative and interdisciplinary contexts. What are the roles of the social scientist, field scientist, and computer scientist in the design and development of new infrastructures? We propose a session to discuss these issues as they relate to the work of EarthCube and its building blocks. We will explore the social aspects of the development of technology, models and systems for supporting scientific knowledge production as well as the outcomes of that science as it moves from the lab into the world. This session will be a panel and group discussion, followed by a description of methods and tools for assessing the social and policy impacts of 'ground-level' and 'high-level' decisions about data management, governance and re-use.

INCREASING THE POTENTIAL USABILITY, ACCESS, AND FUNDING OF OUR GEOLOGICAL COLLECTIONS—PRESENTATION AND OPEN DISCUSSION

Convener: Timm, Sarah (ECLIPSE)

Session Type: Technology Presentation

Description: Geological materials are used every day, in research, as educational materials, as ornamental pieces, and when broken down to their finest components are incorporated into objects we use on a daily basis. For years the accumulation of geological materials in institutional and individual collections has continued without sufficient thought to long-term preservation, access, and usability. Now these collections are at a pivotal point: with geological localities all over the world being shut down, reclaimed, and exhausted, storage limits being exceeded, and limited funding for collections management, physical geological materials are becoming endangered. If action is not taken soon to preserve these collections, then there is great risk of losing them forever.

Due to the technological revolution we have had, the potential for preserving these materials has increased dramatically. Paper labels can only communicate so much information and only to someone nearby the sample. Digital records for specimens can be shared across continents and inform scientists years down to road of analyses preformed today. As part of my master's thesis at Virginia Tech, I developed a database program for geological collections called EGEMS (Electronic Geological Management System). I have spent the majority of the last two years focused on improving access, usability, and funding for geological collections through the use of EGEMS. I have worked with Virginia Tech and the Virginia Museum of Natural History to incorporate existing data and add much more, including data for research materials which could not have easily been done using their previous system. It has been an incredibly rewarding experience but it has had a much smaller impact than it could have given that most academic institutions have geological materials of some kind. I am now working on a web interface called GeoCat to share this data with institutions worldwide. In this session I will discuss the challenges geological collections are facing. I will present some potential solutions including a case example of EGEMS at the Virginia Museum of Natural History. I will demo the functionalities of EGEMS

live, and then open the floor up to a discussion on the access, funding, preservation, and overall future for our collections.

DOMAIN REGISTRIES OF INFORMATION RESOURCES: A HOW-TO FOR YOUR COMMUNITY

Convener: Zaslavsky, Ilya (University of California San Diego)

Co-Conveners: Leslie Hsu (LDEO, Columbia University), Douglas Fils (Consortium for Ocean Leadership), Stephen Richard (Arizona Geological Survey), David Valentine (University of California San Diego), and Kerstin Lehnert (LDEO, Columbia University)

Session Type: Workshop/Hack-a-Thon

Description: Creating domain catalogs of information resources, to support discovery of databases, software, models, portals, vocabularies, and other resources geoscientists used in their work, has been one of the key themes of EarthCube end-user workshops. Several workshops have assembled such initial catalogs. In addition, the EarthCube SEN and C4P Research Coordination Network projects created such catalogs for their domains. Working with the EarthCube CINERGI project, the SEN and C4P projects established a model of how such catalogs can be assembled, published, and community-curated.

This workshop, jointly presented by members of the CINERGI, C4P and SEN teams, will introduce this model, and walk you from assembling a seed catalog, to populating it with standard metadata, to publishing it on the web in a way that supports catalog visualization and community updates of resource information. The goal is to facilitate organization of working groups that would create similar catalogs in additional geosciences domains, and help them get started leveraging our software tools and experience.

In addition, the workshop will include a general introduction to issues and technologies being used to describe catalog harvesting and curation tools being developed as part of the CINERGI project.

Appendix 2: Institutions and Organizations Represented

1. AGU
2. American Geosciences Institute
3. American Meteorological Society
4. Antarctic Glaciological Data Center (AGDC)
5. Arizona Geological Survey
6. Aron Environmental Consulting
7. Boston University
8. Brigham Young University
9. California Institute of Technology
10. case western reserve univesity
11. Chief of Staff Office
12. CIESIN, Columbia University
13. Clifford A Jacobs, Consulting, LLC
14. College of Charleston
15. Columbia University
16. Computing, Data and Software Facility, NCAR/EOL
17. Consortium for Ocean Leadership
18. CSDCO/LacCore - Univ of MN
19. CUAHSI
20. ECLIPSE
21. EDC
22. Element 84
23. Foundation for Earth Science
24. George Mason University
25. Georgia Institute of Technology
26. GFZ German Research Centre for Geosciences
27. GMU/CISC
28. Indiana University Bloomington
29. INSTAAR Univ Colorado
30. iPlant
31. IRIS
32. J&FE
33. JPL
34. Lamont-Doherty Earth Observatory
35. MapStory
36. Marymount University
37. Microsoft
38. Mineralogical Society of America
39. Minnesota Geological Survey
40. National Center for Atmospheric Research

41. National Center for Supercomputing Applications
42. National Science Foundation
43. National Snow and Ice Data Center
44. NCAR
45. NEON Inc.
46. New Jersey Institute of Technology
47. New Media Research Institute
48. NJIT
49. NOAA
50. NSIDC
51. Old Dominion University
52. OPeNDAP
53. Oregon State University
54. Penn State University
55. Princeton University
56. RENCI
57. Research Data Alliance
58. San Diego Supercomputer Center
59. Scripps Institution of Oceanography
60. University of S
61. Semantic Community
62. Sematron, LLC
63. SOCoP
64. Sonoma State University
65. Stroud Water Research Center
66. Texas A&M University
67. The Aerospace Corporation
68. The HDF Group
69. University of Kansas
70. Tufts University
71. U.S. Geological Survey
72. UCAR Unidata
73. UMBC
74. UMD
75. UNAVCO
76. United States Geological Survey
77. Univ at Buffalo
78. University at Albany, SUNY
79. University of Akron
80. University of Alabama in Huntsville
81. University of Arizona
82. University of California, Irvine
83. University of California, Los Angeles
84. University of California, Riverside

85. University of California, Santa Barbara
86. University of Chicago
87. University of Hawaii, Manoa
88. University of Illinois at Chicago
89. University of Maryland
90. University of Michigan
91. University of South Florida
92. University of Southern California/Information Sciences Institute
93. University of Texas at Austin
94. University of Utah
95. University of Washington
96. University of Wisconsin-Madison
97. Virginia Tech
98. Woods Hole Oceanographic Institution
99. Wright State University