2012

EarthCube Governance Roadmap

Documentation, Research, and Recommendations

Version 2.0

August 2012

EarthCube Governance Working Group
presented to the National Science Foundation
8/15/2012

EARTHCUBE GOVERNANCE ROADMAP STEERING COMMITTEE

M. Lee Allison, Chair, Arizona Geological Survey
Tim Ahern, IRIS Data Management
David Arctur, University of Texas at Austin
Jim Bowring, College of Charleston
Gary Crane, Southeastern Universities Research Association
Cecelia DeLuca, NESII/CIRES/NOAA Earth System Research Laboratory
Geoffrey Fox, Indiana University
Carroll Hood, Raytheon
Hannes Leetaru, Illinois Geological Survey
Kerstin Lehnert, Columbia University
Chris MacDermaid, Colorado State University/CIRA
Mohan Ramamurthy, Unidata
Erin Robinson, Federation of Earth Science Information Partners
Ilya Zaslavsky, San Diego SuperComputer Center

Arizona Geological Survey Staff
Genevieve Pearthree
Kim Patten
Debra Winstead

In addition, there were countless members of the EarthCube Governance Forum, EarthCube community, and other interested partners that contributed to this report.

PREFACE

EarthCube is a process established to transform the conduct of research through the development of community-guided cyberinfrastructure. Specifically, EarthCube is attempting to integrate data and information knowledge management across the Geosciences; providing capabilities that permit scientists to conduct research in a more productive way, integrating their work with data from other agencies as well as international partners. Thus, EarthCube aims to create a knowledge management system and infrastructure that integrates all Earth system and human dimensions data in an open, transparent, and inclusive manner. EarthCube requires broad community participation in concept, framework, and implementation, and must not be hindered by rigid preconceptions.

We were tasked with creating a roadmap for the future governance of EarthCube. This paper provides three complete units for accomplishing this. Part 1 – the Roadmap – provides action items and a timeline for developing and implementing a governance framework. Part 2 – the Roadmap Documentation – provides insight into the ten guidance points presented by the National Science Foundation (NSF) at the onset of this project. Part 3 – provides a complete summary of the research and community outreach conducted to produce parts 1 and 2.

This material is based upon work supported by the National Science Foundation
Under Grant No. 1238951 and 0753154.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

ROADMAP

EXECUTIVE SUMMARY

INTRODUCTION

In order to adequately absorb the concepts and drivers behind the motivations and suggestions in this roadmap and subsequent documentation, we first must define the concept of Governance. For our purposes, we the crafted following definition based on governance research and community input:

> ***Governance*** *refers to the processes, structure and organizational elements that determine, within an organization or system of organizations, how power is exercised, how stakeholders have their say, how decisions are made, and how decision makers are held accountable.*

A key question the Governance Working Group struggled with is whether something as complex as EarthCube can be governed by one single model or by one encompassing entity. The unstated presumption entering the governance roadmap-writing process was that there would be a single governing entity for EarthCube and that such an approach is inherently necessary to proceed. There is long standing concern among some in the research community, however, whether something this important should be managed by a single entity. We have learned, from historical infrastructure case studies, background research on governance theory and models, and from community feedback during this roadmap process, that many large-scale, complex infrastructures, including the Internet, have no central control, administration, or management. No national infrastructure that we examined is governed by a single entity, let alone a single governance archetype. While this is precedence, we feel that through an aggressive community engagement process, EarthCube can do better, maximizing upon lessons learned.

Six months of governance research and engagement of potential EarthCube stakeholders provided a venue to begin to tackle these challenging and complex questions. Although limited time and resources constrained the scale of our research review and community engagement, they did provide the foundation to forge a viable path forward in setting of a governance framework for EarthCube. Thus, this roadmap represents a compilation of the governance information we have gathered to-date, with governance recommendations based on what we've learned throughout the roadmap-writing process. As further research occurs, these recommendations may change.

Recommendations

We feel the roadmap process must accommodate an organic governance system or system of systems. EarthCube may begin with a single governing entity or structure, but likely will evolve into a collective of governing bodies as warranted, in order to be successful. Additionally, we believe EarthCube governance will need to be flexible to accommodate a range of governance models that meet the needs of each of the sub-discipline communities and achieve the goals of the National Science Foundation (NSF), have the ability to address foreseen and unforeseen challenges, and be able to adequately address requirements identified by the community constituents.

We recommend a successor governance body be established to evaluate the single entity vs. system of systems approaches as part of the implementation of the Governance Roadmap for EarthCube. These should be considered for both short-term (i.e., design and build) and long-term (i.e., adoption and operation) goals.

We recommend a detailed timeline with action items and target dates, presented both in in the Milestones and Tasks section immediately following the Executive Summary:

1. August 31, 2012: Determine the Governance Framework to meet community needs and NSF goals for successful cyberinfrastructure.
2. August 31, 2012: Determine the stakeholder community and identify initial governance committee for engaging the community for input on the governance framework developed in Step 1.
3. Date TBD: Establish Terms of Reference for EarthCube Governance.
4. Date TBD: Implement the suggested EarthCube Governance Terms of Reference.

It is the position of this Governance Working Group that the governance roadmap should be executed expeditiously and be completed in large part before the execution of other roadmaps. Thus, this roadmap should be completed as soon as feasible after the second NSF EarthCube Charrette in June, 2012, in order to provide groups and individuals focusing on more technical aspects of EarthCube with guidance to make informed decisions about the management environment in which they will operate.

NSF GUIDANCE POINTS

NSF provided each Working Group and Concept Team with 10 guidance points with the goal to expand and enhance each group's ability to write an actionable roadmap to move EarthCube forward. These 10 guidance points are explored extensively in the Roadmap Documentation Section. A brief summary of each guidance point is provided here.

1.0 Purpose

The overarching goals of EarthCube are to build a unified, adaptive, and scalable cyberinfrastructure framework for enabling transformative advances in geosciences research and education, thereby realizing the objectives articulated in the NSF Geo Vision report.[1] In the process, EarthCube aims to create a knowledge management system and infrastructure that integrates all Earth system and human-dimensions data (data related to human uses of the Earth) in an open, transparent, and inclusive manner.

Developing a viable organizational and governance structure for any organization is a formidable challenge. Creating an effective structure for multi-disciplined, distributed, virtual collection of scientists, investigators, technologists, system operators, entrepreneurs, and administrators can be nearly impossible unless great care is taken to ensure that the proposed solution is flexible and responsive to meet participant's needs and institutional goals.

We believe that there is general agreement that "effective governance for EarthCube will:

1. Actively engage its diverse users

[1]National Science Foundation, Advisory Committee for Geosciences, "GEO Vision Report." October 2009.

2. Provide leadership and oversight to forge close cooperation, coordination, and collaboration among distributed development activities and the principal EarthCube groups
3. Facilitate alignment of funding program plans and priorities with the needs of the community
4. Guide the successful execution of the EarthCube mission, meeting stakeholder obligations."[2]

To be effective, the governance framework the community adopts is likely to be for a *system* of governance (a matrix of mechanisms for different elements and groups) that accommodates different practices and requirements among different elements of a large and diverse community. The governance roadmap also allows for a variety of mechanisms for how the governance mechanisms are chosen and implemented.

2.0 Communication

Community input on issues to consider and materials to review guided the governance roadmap-writing process. Weekly Governance Steering Committee calls, plenary sessions, outreach webinars to industry, government agency, and international representatives, extensive use of the EarthCube Ning site, and official liaisons to each of the EarthCube Working Groups and Concept Teams formed the backbone of the communications to-date. The liaisons coordinated between the more technically focused groups and the Governance Working Group to help ensure unique governance issues and needs from each segment of EarthCube were identified and addressed.

An additional Governance Forum, composed of individuals and organizations forming a larger, broader constituency that had demonstrated their interest and participation in governance issues, was formed. One of the goals of organizing the Forum was to ensure representation from each of the Geoscience and closely related domains (e.g. Earth, Atmospheres, Ocean, Environment and Ecology), IT sectors, academia, agencies, and industry, involved in determining agendas, setting goals, and having an active role in formulating the roadmap.

Future community engagement and communication will build on processes developed during the writing of the EarthCube Governance Roadmap. A community engagement program to gather EarthCube governance requirements and vet initial governance framework recommendations may consist of:

1. Gathering requirements for the earthcube.org website
2. Engaging the geosciences, atmosphere, and ocean communities
3. Engaging the computer science, IT and software communities
4. Mapping the EarthCube stakeholder community
5. Developing materials to market EarthCube to potential new users
6. Holding webinars for target communities of practice
7. Engaging existing EarthCube Working Groups and Concept Teams
8. Publishing EarthCube articles in trade journals

[2]Mohan Ramamurthy, "Unidata Governance: A Quarter Century of Experience," National Science Foundation EarthCube White Paper: Governance Category, 2011, 1.

9. Engaging domain and IT communities through a stakeholder survey
10. Engaging the community through social media
11. Attending conferences of opportunity, such as ESIP and AGU, to engage potential EarthCube stakeholders
12. Employing public listservs, wikis, and the EarthCube Ning site to communicate with EarthCube stakeholders and provide an additional forum for community input

3.0 Challenges

The challenges we considered in writing this roadmap were not just to creating the governance roadmap per se, but also to the role and impacts of a governance process and system on the overall viability and success of EarthCube as a community system. Challenges to the roadmapping process are inherent given the limited time frame, including:

1. Comprehensive background research review of governance topics from the domain sciences, IT, and social sciences is not yet complete.
2. We identified many governance models, but have not been able to fully evaluate them.
3. Further work is needed to evaluate the pros and cons of different models and determine which may be suitable for EarthCube.
4. Our knowledge of the other EarthCube Working Group and Concept Team governance issues and needs is not yet complete.
5. We have yet to fully engage the broader Earth, information, and IT science communities. Thus our knowledge of their governance needs is limited.
6. There is limited information available about problems and failures of past projects that we can incorporate as things to avoid.

Challenges to the viability of EarthCube were generated by community feedback and the governance research review. We divided them into:

1. <u>Conceptual and procedural challenges:</u>[3] Time (short-term funding decisions versus the long-term time-scale needed for infrastructures to grow); Scale (choices between worldwide interoperability and local optimization); Agency (how to navigate planned versus emergent change), intellectual property rights, infrastructure winners and losers, agreement on data storage, preservation, curation policies and procedures, incentives to share data and data sharing policies, and trust between data generators and data users.
2. <u>Social and cultural challenges:</u> Motivations and incentives, self-selected or closely-held leadership, levels of participation, types of organizations, and collaboration among domain and IT specialists)
3. <u>Technical challenges:</u> From governance use cases.
4. <u>Trends and drivers:</u> Federal government initiatives, cloud computing.

4.0 Requirements

To continue forward, we recommend building upon the process of community engagement and research review begun as a cornerstone of the Governance Roadmap process to identify and

[3] Paul Edwards, Steven Jackson, Geoffrey Bowker, and Cory Knobel, "Understanding Infrastructure: Dynamics, Tensions, and Design - Report of a Workshop on "History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures," 2007, 24-33.

characterize the governance components of cyberinfrastructure. Community engagement is expected to occur in four steps (for a full description and graphic showing the progression of engagement see Roadmap Documentation Section 4: Requirements):

1. Identify cyberinfrastructure components of EarthCube.
2. Identify the cyberinfrastructure components' organizational paradigms and governance needs.
3. Identify the interaction among and between cyberinfrastructure components and systems within EarthCube.
4. Identify the interactions between cyberinfrastructure components within and outside of EarthCube, and the needs of EarthCube consumers (including those composing the "long tail" of science).

The process of community engagement described above, in addition to extensive background governance research review (see Appendix 1), generated a list of initial requirements for EarthCube governance. EarthCube governance framework requirements identified thus far include:

1. Best practices and lessons learned
2. Barriers to participation
3. Collaboration
4. Communication
5. Community
6. Culture
7. Decision-making
8. Design
9. Education and outreach
10. Evaluation
11. Evolution
12. Goals and objectives
13. Governance as a system and multiple levels of governance
14. Integration and interoperability
15. Leadership
16. Legal issues
17. Metrics
18. Reproducible science
19. Risk Management
20. Scale
21. Sustainability
22. Temporal aspects
23. Trust
24. Vision

These requirements will be expanded upon and categorized in the Governance Framework to be presented to NSF and the EarthCube community by the end of August 2012 (Step 1 of the Milestones and Tasks).

5.0 Status

Drawing on the experience of IT governance, principally in the business environment but in other organizations, as many as seven different decision-making processes may be employed within an organization, ranging from monarchy to federal to anarchy (or crowdsourcing) and variations in between. Instead of a single monolithic model deployed across all elements, governance can be viewed as a system composed of many (potentially overlapping) models used as needed to generate information and make the types of decisions to ensure effective management and use of IT. It is necessary to resolve social, cultural, and organizational issues in tandem with IT services in order to create a strong and lasting infrastructure.

In addition, social science research identifies not only the options for governance but also the process for adoption and implementation of governance and organizational structure in the technical and scientific magisteria.

We also drew on research on the evolution of all infrastructures, which identify common characteristics across historical infrastructure development. Infrastructures start out with interplay among largely independent and even disparate elements that grow, link with other elements, and create and manage interactions among the components. Such is the current path for cyberinfrastructure.

6.0 Solutions

The objective of the EarthCube governance framework is to guide the development of infrastructure that will become ubiquitous and effectively invisible to users. Thus, the currently preferred governance framework for distributed cyberinfrastructure like EarthCube is a system of systems approach, responsible for managing the dependent, independent, and interdependent components that form complex infrastructures such as EarthCube. In this way, different governance models are employed to govern different aspects of the cyberinfrastructure. Thus, a single, one-size-fits-all governance model may not be appropriate for an entity as large and complex as EarthCube is envisioned to be. However, these concepts are designed to encourage critical thinking on governance at this point and the theory may change as more of the community is engaged

Of all the types of infrastructure we examined for this study, the Internet may be the closest analog to EarthCube. The Internet has no centralized governance in either technological implementation or policies for access and usage; each constituent network sets its own standards. There is no central control, administration, or management of the Internet, therefore it is misleading to use the term 'Internet Governance' when 'The Internet' is not a single entity to govern. Instead, several organizations work together in a relatively well structured and roughly democratic environment to collectively participate in the research, development, and management of the Internet. This is closer to how EarthCube is currently governed, but there is no agreement that this is the best or even a viable model for moving on to the next phase.

7.0 Process

Despite an intimidating list of challenges that we compiled, there is a growing realization that the initial roadmap is not required to solve all of them. In fact, the June 2012 NSF Charrette provided an ideal venue to address and try to resolve many of them in consultation with the other community and concept groups. We also have scheduled post-Charrette activities to revise the roadmap based on what was learned and discussed there. We envision a successor body to the Governance Work Group, generically named the Interim Governance Committee that will continue to handle these challenges as part of the implementation process. The Milestones and Tasks immediately following the Executive Summary describes in more detail the subsequent process and steps.

8.0 Timeline

We recommend a detailed timeline with action items and target dates:

1. <u>August 31, 2012:</u> Determine the Governance Framework to meet community needs and NSF goals for successful cyberinfrastructure.

2. August 31, 2012: Determine the stakeholder community and identify initial governance committee for engaging the community for input on the governance framework developed in Step 1.
3. Date TBD: Establish Terms of Reference for EarthCube Governance.
4. Date TBD: Implement the suggested EarthCube Governance Terms of Reference.

9.0 Management

Governance roadmap implementation begins with the formation of an Interim Governance Committee to determine the initial EarthCube stakeholder community, analyze potential governance models, and develop a framework for governance. It is important to examine the potential roles and responsibilities, support processes, decision-making processes, and contribution processes that may compose EarthCube governance.

10.0 Risks

Risks regarding the rapid development of a roadmap for EarthCube governance include: 1) radical departure from the traditional NSF process, 2) group think and limiting perceptions of the Governance Steering Committee members, 3) inability to decide upon the scope and structure of the governance process, and 4) short-term thinking and development of a short-term governance framework.

Risks in meeting EarthCube goals include: 1) a lack of past examples and lessons learned, 2) lack of contingency plans, 3) interweaving architectures, 4) tyranny of the major players over the small or distributed users, 5) development of a governance approach that proves unwieldy or ineffective, and 6) sustainability and evolution of governance. Additional risks include 1) uncertainty of institutional commitment and support, 2) commitment of the potential user community, and 3) risks identified by other EarthCube roadmaps.

MILESTONES AND TASKS

Goal: To provide a roadmap such that an optimal implementation – agnostic governance framework can be put in place to meet the community needs and achieve the goals of EarthCube.

<u>Step 1:</u> Determine the appropriate Governance Framework to meet community needs and NSF goals for successful cyberinfrastructure

TIME: begin with Charrette, complete by August 31, 2012

DELIVERABLE: Document reporting findings and decisions leading to selection of Initial Governance Framework

1. What Governance models may be best suited for EarthCube?
    a. How will competitive solicitations be enabled, if required?
    b. How will decisions be made?
    c. How will the governance framework be held accountable?
    d. What are the relationships of governance elements to NSF?

2. What are the scopes of authority of governance within and among elements of EarthCube?
    a. Who is in charge of creating and updating EarthCube milestones and determining how they will be achieved? [Key Question for Charrette]
    b. A governance system provides overall steering and oversight for
        i. Leadership
        ii. Adjudication
        iii. Policy development (Science Priorities)
        iv. Intellectual property right protections
        v. Data publishing & data citation
        vi. Model publishing & model citation
        vii. Other use cases
        viii. Sustainability
    c. Spending authority, pass-through funding coordination
    d. Operations
    e. Outreach and communications
    f. Proactive coordination
    g. Architecture maintenance and systems integration support (Technical Priorities)
        i. Standards selection processes supported by:
            1. Prototyping and risk reduction
            2. Testbeds
        ii. Life-cycle data management
            1. Certification of roles
        iii. Security architecture
            1. Authentication
            2. Identity Management
            3. Single sign-on
            4. Threat Mitigation
            5. Policy decision and enforcement
        iv. Registry Management
        v. Other use cases

3. How will the community be involved?
    a. Management and organization - within each community-driven governance element, there could be
        i. Board of Directors
        ii. Advisory Committee structure
        iii. Science Advisory Committee
        iv. Technical Advisory Committee
        v. Program Committees (e.g., Finance, Nominations, Policy, etc.)
        vi. Functional Committees (e.g., Workflow, Data Management, etc.)
        vii. Identification of desired communities of practice. (E.g. Community of practice for outreach and education about current standards, conventions, and best practices for data models, model frameworks, and software engineering. This would be cross-cutting among all science domains, and use social networking tools and outreach.)
    b. End-user engagement & community ownership
        i. Decision and policy makers
        ii. Existing data centers
        iii. Long-tail scientists

4. How will EarthCube governance functions be supported? Any or all of the following methods?
    a. Grants and Contracts Funded
    b. Volunteer
    c. Membership Fees
    d. Public/Private Partnerships
    e. International Initiatives
    f. Provision of Services
5. How will the EarthCube Governance structure manage its own evolution?
    a. What processes are there to recognize new needs as they emerge?
    b. Governance structure should be dynamic

<u>Step 2:</u> Determine the stakeholder community and identify initial governance committee for engaging the community for input on the governance framework developed in Step 1.

TIME: begin with Charrette, complete engagement plan by August 31, 2012, complete initial engagement by beginning of 2013

DELIVERABLE: Document reporting findings and decisions, identifying who is involved

1. Much of this is being done through the current EarthCube community building process that NSF has initiated
    a. Coordination with NSF Concept Awards and Working Groups
    b. Coordination with Focus Groups (e.g. Industry, International, Federal Agencies, Education, etc.)

2. Identify an initial representative group to act for the larger community (EarthCube Governance Committee) to implement the Roadmap
    a. What size should initial group be?

b. What communities need representation?
c. What role does NSF play in this?
d. Determine how this initial group is selected.
e. Identify linkages to other existing NSF cyberinfrastructure facilities
f. Leadership -
    i. Who is in charge of making decisions?
    ii. What decisions need to be made?
    iii. Does the individual/group in charge of making decisions change depending on what type of decision is being made?
    iv. Who has the authority to resolve disputes?
    v. What mechanisms are in place to allow for new leadership to emerge if/when necessary?

3. Craft a community engagement program to gather EarthCube governance requirements and vet the initial Governance Framework developed in Step 1
    a. Gather requirements for the earthcube.org website
    b. Engage the geosciences, atmosphere, and ocean communities
    c. Engage the computer science, IT and software communities
    d. Map the EarthCube stakeholder community
    e. Develop materials to market EarthCube to potential new users
    f. Hold webinars for target communities of practice
    g. Engage existing EarthCube Working Groups and Concept Teams
    h. Publish EarthCube articles in trade journals
    i. Engage domain and IT communities through a stakeholder survey
    j. Engage the community through social media
    k. Attend conferences of opportunity, such as ESIP and AGU, to engage potential EarthCube stakeholders
    l. Employ public listservs, wikis, and the EarthCube Ning site to communicate with EarthCube stakeholders and provide an additional forum for community input

<u>Step 3:</u> Establish Terms of Reference for EarthCube Governance

TIME: Date To-Be-Determined

DELIVERABLE: Initial Terms of Reference for EC Governance Entity that covers the following items

1. EarthCube Governance Committee tasked with drafting an EarthCube Governance Terms of Reference and other documentation that define and codify the Governance characteristics agreed to in the previous steps. This will include but is not limited to:
    a. A process for creating a permanent EarthCube Governance Framework
    b. Decision making structure and leadership
    c. Determination of issues regarding membership, representation, voting, planning
    d. Address issues of evolving EarthCube goals
        i. Who is in charge of creating and updating EarthCube goals? [Key Question for Charrette]
        ii. How does this relate to governance framework?
    e. A transition plan from initial to permanent framework

f. Participation policy - what does it mean to be a participating EarthCube community member, developer, contributor, etc.?
g. Process to gather and integrate feedback on proposed EarthCube governance structure from broader community

2. Revise proposed EarthCube Terms of Reference based on feedback

3. Finalize EarthCube Terms of Reference and seek approval of the EarthCube community

Step 4: Implement the suggested EarthCube Governance Terms of Reference

TIME: Date To-Be-Determined

DELIVERABLE: Document reporting decisions; workshop series

GANTT CHART

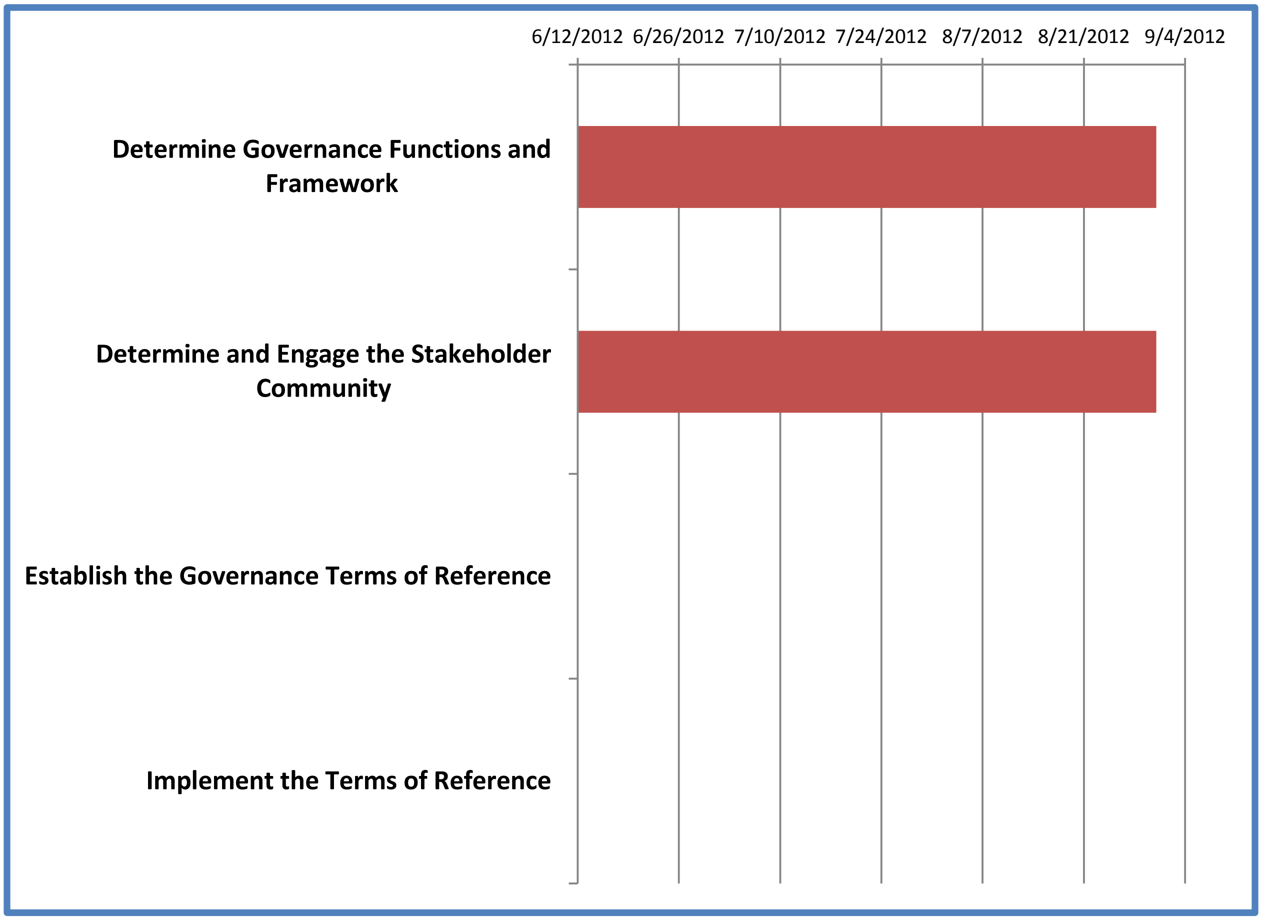

FIGURE 1: GANTT CHART FOR IMPLEMENTATION OF EARTHCUBE GOVERNANCE MILESTONES AND TASKS.

ROADMAP DOCUMENTATION

The section provides further insight into the recommendations made in the actionable roadmap introduced in the previous section. We use the NSF Roadmap Guidance, or 10 points, to outline our research findings.

1.0 PURPOSE

1.1 INTRODUCTION

"All parties—scientists, societies, and funders—must play a role in creating governance around controlled vocabularies, taxonomies, and ontologies that can be used in scientific applications to ensure the currency and evolution of knowledge encoded in semantics"[4]

"To state it baldly, the paradigm that needs destruction is the idea that we as scientists exist as un-networked individuals"[5]

"Funding agencies must increasingly target the building of communities of practice, with emphasis on the types of interdisciplinary teams of researchers and practitioners that are needed to advance and sustain semantic eScience efforts"[6]

"We increasingly need to embrace e-Research techniques and use the internet not only to access nationally distributed datasets, instruments and compute infrastructure, but also to build online, 'virtual' communities of globally dispersed researchers."[7]

"The alignment of science supply and demand in the context of continuing scientific uncertainty will depend on seeking out new relationships, overcoming language and cultural barriers to enable collaboration, and merging models and data to evaluate scenarios.… Capturing the important elements of data preservation, collaboration, provenance, and accountability will require new approaches in the highly distributed, data-intensive research community.[8]

"Communities of Practice are groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis"[9]

---

[4] Peter Fox and James Hendler, "Semantic Escience: Encoding Meaning in Next Generation Digitally-Enhance Science," in *The Fourth Paradigm*, ed. Tony Hey, Steward Tansley, and Kristin Tolle (Redmond, WA: Microsoft Research, 2009), 147-152.

[5] John Wilbanks, "I Have Seen the Paradigm Shift, and It Is Us." ," in *The Fourth Paradigm*, ed, Tony Hey Steward Tansley, and Kristin Tolle (Redmond, WA: Microsoft Research, 2009), 209-214.

[6]Fox & Hendler, "Semantic Escience," 147-152.

[7]"Earth Sciences and Beyond…Why Build an Integrative Earth Science e-Research Infrastructure?" AuScope GRID, last modified August 2, 2011, http://www.auscope.org.au/site/grid.php.

[8]Mark Abott, "A New Path for Science?" in *The Fourth Paradigm*, ed. Tony Hey, Stewart Tansley, and Kristin Tolle (Redmond, Washington: Microsoft Research, 2009), 114-116.

[9]Etienne Wenger, Richard McDermott, and William Snyder, *Cultivating Communities of Practice,* (Boston, MA: Harvard Business School Press, 2002).

The overarching goals of EarthCube are to build a unified, adaptive, and scalable cyberinfrastructure for enabling transformative advances in geosciences research and education and realizing the vision articulated in the NSF Geo Vision report.[10] In the process, EarthCube aims to create a knowledge management system and infrastructure that integrates all Earth system and human dimensions data in an open, transparent, and inclusive manner.

A specific goal of EarthCube is to seek agreement and initiate an ongoing process for the entire geoscience community and its stakeholders through coordinated approaches to cyberinfrastructure, which is the very definition of sustainable cyberinfrastructure. With enhanced access to large, federated data and service networks, and the broader geoscience community, including academia and the private sector, scientists may use the synoptic data resources efficiently and integrate them with their own data and research.

The meshing of the domain sciences and the IT communities is fundamental to the success of the EarthCube vision. This integration is also emblematic of the underlying duality of science and infrastructure. The former depends on a culture of experimentation, flexibility, rapid change, and innovation, or in effect, a certain amount of anarchy. The latter requires continuity, certainty, and deliberative processes for incorporating change and cost effectiveness, and is inherently more centralized. The infrastructure should be largely transparent to the users of the system, but the coordination of the two environments is critical. Thus, the governance of scientific cyberinfrastructure requires mechanisms to identify and manage synergy and autonomy among its disparate components.

The project of building a user-driven framework that addresses the computational, data, software, knowledge management, and user service needs of the diverse earth system science community is enormous. However, the biggest challenges in past infrastructure development have been sociological and organizational rather than technical.[11] Therefore EarthCube's governance must recognize and address the need to bridge cultural differences between disparate disciplines and developers, as well as user communities, bringing all stakeholders together to create system-wide solutions to community problems.[12]

Keeping this in mind, we believe that there is general agreement that "effective governance for EarthCube will:

- "Actively engage its diverse users
- Provide leadership and oversight to forge close cooperation, coordination, and collaboration among distributed development activities and the principal EarthCube groups
- Facilitate alignment of funding program plans and priorities with the needs of the community
- Help the successful execution of the EarthCube mission, meeting stakeholder obligations"[13]

[10] National Science Foundation, Advisory Committee for Geosciences, "GEO Vision Report."
[11] Edwards et al., "Understanding Infrastructure."
[12] Mohan Ramamurthy, "Unidata Governance."
[13] Mohan Ramamurthy, "Unidata Governance," 1.

We recognize that there exist models and implementations of governance within the scientific cyberinfrastructure arena, notably, the governance structures for Open Geospatial Consortium (OGC), TeraGrid, Open Grid Forum, Earth Science Information Partners (ESIP) Federation, CUAHSI-HIS, IRIS, Unidata, GENI, and SURAGrid, among many others. We also recognize that there are governance models for more complex systems system, such as the U.S. electric grid and the internet, that provide useful examples of how infrastructure, and cyberinfrastructure, can be governed. All of these models may provide useful insights regarding the development of an optimal governance for EarthCube.

1.2 BACKGROUND AND HISTORY

Cyberinfrastructure is being built by individuals, teams, groups, and communities. One example of this is the November 2011 NSF EarthCube Charrette, which demonstrated pervasive agreement on the basic attributes of EarthCube, and the need to coordinate the diverse array of activities already underway in the geosciences, related domain sciences, and information technology arenas.

There are a number of workshops and reports over the past decade that serve as references for the Governance roadmap process. Summaries of these reports are included in Appendix 1: Research Review.

1.3 WHAT IS MEANT BY "GOVERNANCE" AND WHY DO WE NEED IT?

Governance means different things to different groups. EarthCube governance must take all of these definitions into account.

Governance Definitions

Those involved in standards and procedures view it as a way to define relationships between system partners, agreements, contracts and enforcement mechanisms, determine lifecycle management of system components (data, services, catalogs, identifiers, vocabularies, application clients, etc.), specify authority, performance expectations, and risk management strategies, enabling cross-disciplinary and cross-jurisdictional research designs, and formulation of policies needed to orchestrate consensus-building and regulate potential conflicts at the boundaries of subsystems.[14]

The EarthSystem Commodity Governance Project defines governance as "align[ing] an organization's practices and procedures with its goals, purposes, and values. Definitions vary, but in general governance involves overseeing, steering, and articulating organizational norms and processes (as opposed to managerial activities such as detailed planning and allocation of effort). Styles of governance range from authoritarian to communalist to anarchical, each with advantages and drawbacks."[15]

[14] Ilya Zaslavsky, Mark Williams, Anthony Aufdenkampe, Kerstin Lehnert, Emilio Mayorga, and Jeff Horsburgh, "Data Infrastructure for the Critical Zone Observatories (CZOData): an EarthCube Design Prototype," (National Science Foundation EarthCube White Paper: Designs Category), 2011.
[15] "Governance," EarthSystem Commodity Governance Project, last modified 2012, http://earthsystemcog.org/projects/cog/governance_object

Massachusetts Institute of Technology Sloan School of Management and Center for Information Systems Research (CISR) senior research scientist Peter Weill and director Jeanne Ross define IT governance as, "Specifying the decision rights and accountability framework to encourage desirable behavior in using IT. IT governance is not about making specific IT decisions–management does that–but rather determines who systematically makes and contributes to those decisions."[16]

The Organization for Economic Cooperation and Development (OECD) defines governance as, "Providing the structure for determining organizational objectives and monitoring performance to ensure that objectives are attained."[17]

In open source software projects, the "governance model describes the roles that project participants can take on and the process for decision making within the project. In addition, it describes the ground rules for participation in the project and the processes for communicating and sharing within the project team and community."[18]

Additionally, the recent United States Geoscience Information Network (USGIN) strategic plan noted the need for governance to include education, outreach, communication, and marketing to other participants, funders, users, and students, as well as sustainability.[19] We believe these components of governance should also be included in EarthCube.

Governance Implementation

Governance of IT components in larger organizations draw from the business environment,[20] in which different governance models are applied for different functions. The choice of which model to apply for which decision area depends on the organization's strategies, structure, and goals. There is no single matrix that works, or works best, for all groups or even for similar groups. Each decision process may be best enabled by a distinct governance model within a larger governance system or there may be a number of archetypes that work for a decision-making process but which may be chosen for cultural, social, or historical reasons.

Thus, to be effective, the governance of EarthCube that the community adopts is likely to be for a *system* of governance (a matrix of mechanisms for different elements and groups) that accommodates different practices and requirements among different elements of a large and diverse community. The governance roadmap also allows for a variety of mechanisms to determine how the governance framework for EarthCube is chosen and implemented.

---

[16] Peter Weill and Jeanne Ross, *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results,* (Boston, MA: Harvard Business School Press, 2004), 2.
[17] Organization for Economic Cooperation and Development, Directorate for Financial, Fiscal and Enterprise Affairs, *OECD Principles of Corporate Governance*, (April 1999), 5, 219, quoted in Weill and Ross, *IT Governance*, 4-5
[18] Ross Gardler and Gabriel Hanganu, "Governance Models," Open Source Software Watch, last modified February 14, 2012, http://www.oss-watch.ac.uk/resources/governanceModels.xml.
[19]Stephen Richard and Vivian Hutchison, "Recommendations for the Future of the U.S. Geoscience Information Network," (Arizona Geological Survey, 2011), 11p.
[20] Weill and Ross, *IT Governance*.

A short summary of governance definitions, archetypes/models, components of effective governance, and lessons learned from past cyberinfrastructure projects is presented in Section 5.0 Status, and an expanded version may be found in Appendix 1: Research Review.

1.4 COMMUNITIES TO BE SERVED

The communities to be served by EarthCube governance include the current EarthCube participants as well as all organizations currently under the NSF Geosciences Directorate (Atmospheric and Geospace Sciences, Earth Sciences, and Ocean Sciences) and Office of Cyberinfrastructure. However, EarthCube governance must reach beyond these current participants, accommodating those currently involved in the 'long-tail' of science or data integration as well as those involved in like-minded (or data intensive) communities external to NSF, including Federal Agencies, such as NASA, NOAA, and the Department of Defense; State Agencies; or academic communities. Similarly the EarthCube framework must include large-scale national data initiatives such as the White House's BIGDATA and Open Data initiatives. Finally, in order to be all-inclusive, EarthCube governance must consider the role of industry participants as well as international organizations, students and policy makers.

1.5 IMPROVED CAPABILITIES DUE TO GOVERNANCE OUTCOMES

"… Governance can be messy. Governance fosters debate, negotiation, constructive disagreement, mutual education, and often frustration. The process is messy, but good governance arrangements enable individuals representing an enterprise's conflicting goals to reconcile their views to the enterprise's benefit."[21]

Effective governance offers any number of improved capabilities, including but not limited to:

- Improving communication
- Promoting collaboration
- Helping to create and maintain overall goals
- Ensuring daily practices are aligned towards overall goals
- Anticipating and mitigating risk
- Providing mechanisms to respond to changing community needs and emerging technologies
- Evolving as to adapt to emerging needs and technologies
- Defining roles and responsibilities
- Defining leadership positions and how these positions may be filled
- Reconciling conflicts and arbitrating disputes
- Defining clear processes for making decisions

By providing a framework that promotes communication, facilitation and collaboration, outcomes of the EarthCube Governance Roadmap activities should include moving the emerging field of geoinformatics forward; thus, exhibiting the positive implications of EarthCube to the broader community of researchers, educators and engineers. Huge effort has been expended developing

[21] Weill and Ross, *IT Governance*, 85.

data resources and applications across all the domain sciences, but these have typically been developed in stand-alone project environments and all too often disappear when project funding ends. Thus an important component of EarthCube governance will be to incorporate best practices and lessons learned from past projects, thereby anticipating and mitigating past problems, and benefiting the geoscience community as whole by helping to create a more cohesive and sustainable community-based infrastructure.

The geoscience community will also benefit from a more persistent and standardized framework for implementing and deploying such services, such that information tools and data can be migrated from development environments to maintenance environments when projects are complete. Evolution of standards and the knowledge base for their application and extension in practice to make interoperable, loosely coupled, and transportable data and applications will be fostered by the community.

2.0 COMMUNICATION

2.1 COMMUNICATION TO-DATE

Communication and feedback from across the EarthCube stakeholder community were critical in the evolution of this roadmap and supporting documentation. We upheld a commitment to open and transparent processes from the very beginning. Using the EarthCube Ning site, we advertised all meetings, events, and webinars indicating that they were open to public attendance and participation. We provided virtual capabilities for each of the events and recorded each of the face-to-face meetings, plenary sessions, and breakout sessions with allied fields and agencies, placing them on the EarthCube Vimeo site. We also employed a public listserv to communicate within the Governance Steering Committee.

We formed a Governance Forum, composed of individuals and organizations forming a larger, broader constituency that had demonstrated their interest and participation in governance issues, and a representative from each of the EarthCube Community Groups and Concept Award Teams. The latter served as liaisons between the more technically focused groups and the Governance Working Group to help ensure unique governance issues and needs from each segment of EarthCube were identified and addressed. One of the goals of organizing the Forum was to ensure representation from each of the Geoscience and closely related domains (e.g. Earth, Atmospheres, Ocean, Environment and Ecology), IT sectors, academia, agencies, and industry, involved in determining agendas, setting goals, and having an active role in formulating the roadmap. We communicated with Forum members via an additional public listserv.

Community feedback was collected through the EarthCube Ning site and through email contact, although a series of virtual meetings and workshops hosted by the Governance Steering Committee formed the backbone of community input gathered thus far. The virtual workshops were held with different working groups and stakeholders, covering a range of topics, and began to tackle the complexity, scope, and diversity of issues regarding governance of EarthCube. These workshops were exceptional in providing the initial community feedback and basis required to author the Governance Roadmap.

A complete and detailed collection of community feedback and input, in addition to a list of meetings and EarthCube liaisons, is contained in Appendix 4. Below we summarize the input that we have received through June 7, 2012.

2.2 END-USER RESOURCE RECOMMENDATIONS

Community input and feedback on governance produced a number of important and complex questions that EarthCube governance should consider. Community feedback provided us with a list of recommended resources for review, as well as topical issues such as barriers to participation, collaboration, funding, evolution, global engagement, and more. Each is described below.

Recommended Resources

- The Art of the Community, by Jono Bacon[22]
- Double Loop Organizations Reference List, accessed at: **http://cybersocialstructure.org/2012/04/23/double-loop-governance-references/**[23]
- Workforce (Education, Outreach and Training; Underrepresented Populations), accessed at **http://www.nsf.gov/od/oci/taskforces/TaskForceReport_Learning.pdf**[24]
- EarthCube as a Double Loop Organization, accessed at: **http://cybersocialstructure.org/2012/04/04/earth-cube-as-a-double-loop-organization/**[25]
- Information Technology Standardization: Theory, Process & Organizations by C. Cargill (1989)[26]
- Emergence: The connected lives of ants, brains, cities, and software by S. Johnson, 2002[27]
- Workshop on Working towards a National Geoinformatics Community (NGC) USGS Denver Federal Center, Denver, Colorado, September 23-24, 2010[28]
- USGIN Strategic Plan 2011[29]
- Global Environment for Networking Innovations (GENI): Establishing the GENI Project Office (GPO) (GENI/GPO) **Program Solicitation**, NSF 06-601[30]

Community Recommendations from Steering Committee Calls, Plenary Sessions and Focus Groups

- "Designing a Digital Future: Federally Funded Research and Development Networking and Information Technology," President's Council of Advisors on Science and Technology (PCAST) Report
- Earth and Space Science Informatics 2008 Summit
- "Understanding Infrastructure" by Edwards, et.al.
- NSF EarthCube White Papers: Designs Category (25 papers, submitted October 2011)
- NSF Reports on Virtual Organizations[31]
- Open Source Software (OSS) Watch (multiple publications)

---

22 Bruce Caron, EarthCube Governance Group Forum Discussion Post, May 2, 2012, http://earthcube.ning.com/group/governance/forum/topics/research-review-it-governance-how-top-performers-manage-it.

23 Ibid.

24 Geoffrey Fox, EarthCube Governance Group Forum Discussion Post, April 27, 2012, http://earthcube.ning.com/group/governance/forum/topics/call-for-focus-group-participation .

25 Bruce Caron, EarthCube Governance Group Forum Discussion Post, April 4, 2012, http://earthcube.ning.com/group/governance/forum/topics/earth-cube-as-a-double-loop-organization .

26 David Arctur, EarthCube Governance Group Forum Discussion Post, March 30, 2012, http://earthcube.ning.com/group/governance/forum/topics/initial-earthcube-governance-references-and-works-cited-list.

27 Ibid.

28 Steve Richard, EarthCube Governance Group Forum Discussion Post, February 28, 2012, http://earthcube.ning.com/group/governance/forum/topics/national-geoinfomatics-community-report

29 Steve Richard, EarthCube Governance Group Forum Discussion Post, February 28, 2012, http://earthcube.ning.com/group/governance/forum/topics/usgin-strategic-plan-2011.

30 Recommended by Carroll Hood, Focus Group Webinar, "Engaging Industry Partners on Governance," May 31, 2012.

31 Sayeed Choudhury, community discussion comment, "EarthCube Governance Virtual Plenary Session, April 17, 2012

2.3 COMMUNITY DRIVEN ISSUES AND OPPORTUNITIES

Potential EarthCube stakeholders had several recommendations for EarthCube governance. Some of these recommendations are direct quotes; others are paraphrased.

Barriers: When asked about the existence of potential barriers to participation, one individual replied, "Instead of barriers, there are opportunities to promote interoperability mechanisms."[32]

"The principal barrier to participation in EarthCube is that the domain groups (including CUAHSI) continue to have a primary allegiance to specific domains. I am concerned that EarthCube could compromise the primary missions of the domain-specific groups, by siphoning resources away from serving the needs of domain users."[33]

Case Studies: EarthCube should "[Generalize] from specific success stories to develop new general-purpose services."[34]

One individual presented the iRODS data grid governance was presented as a short governance case study. Governance is accomplished through a series of components:[35]

1. "[Mechanisms to determine the] purpose of collaboration (shared collection, shared research initiative)
2. Properties that the collaboration environment will maintain (completeness, integrity, standards, access)
3. Policies that enforce the desired properties (computer actionable rules)
4. Procedures that derive the desired properties (computer executable workflows)
5. State information that tracks results of applying procedures
6. Assessment criteria that verify desired properties have been enforced"

Communication: It is important to consider the role of communication at EarthCube[36]. In particular, a communication plan should be part of the roadmap and be set in place before the June 2012 Charrette.[37]

Community: "There are multiple user communities: Researchers in national scale projects, Researchers in the "long tail" of science, and Students. They all face the same challenges with quantifying context for research data and research results. There is generic infrastructure that enables each community to provide the needed context. One can build either a Student digital library, or a project data management system, or a collaboration environment with a policy-based data management system. The policies and procedures may be different in each environment, but the underlying data, information, and knowledge virtualization mechanisms are the same."[38]

---

[32] Reagan Moore, personal email, May 25, 2012.
[33] Alva Couch, personal email, May 29, 2012.
[34] Ibid.
[35] Reagan Moore, May 25, 2012.
[36] William Michener, community discussion comment, "EarthCube Governance Virtual Plenary Session," April 17, 2012.
[37] Erin Robinson, personal email, June 2, 2012.
[38] Reagan Moore, May 25, 2012.

"I am assuming that the two governance groups… are representative committees of the community, perhaps elected by the community."[39]

Culture and Global Engagement: A member of the international community recommended: "If EarthCube will be expanding to include significant global collaboration, all governance aspects should consider cultural impacts. For instance, if a particular governance area were to adopt the meritocracy governance archetype, due consideration should be given to culturally induced 'silent/unseen work ethic.'"[40]

Community-Based Collection Cycle: "Instead of a data life cycle, the focus is on a community-based collection life cycle. A collection defines the context for research data. Within a research project, the context may be minimal, since the members of the research team have tacit knowledge on how the data were generated, what processes were applied, and the properties of the individual files such as naming convention. When the user community is broadened, say through a data grid that enables other research teams to collaborate on analysis of the data, the tacit knowledge must be made explicit. This requires transition from governance policies for a research collection to governance policies for a data grid (shared collection). The tacit semantics, data formats, and processing algorithms now need to be defined for use by the additional research team. When the collection is made public for use by the entire discipline, the user community broadens again. Now expectations for the collection context must be mapped to the community consensus on semantics, descriptive metadata, and data formats. When the collection is preserved in a reference collection, the user community broadens again to include future researchers. The context must now include the information and knowledge that defined the original research question, the governance that controlled the approach to solving the research question, and the metrics for success. Within a community-based collection life cycle, the governance policies evolve to track the consensus of each broader user community. This is a direct measure of the NSF mandate to track broadened impact of research results."[41]

Decision-Making: "I strongly urge a consensus-based model. However the consensus is relative to the community that is exploring a research topic." [42]

"Assertions about the community consensus are turned into properties that each object in a collection should possess. Policies are defined to ensure the desired properties are enforced. The policies are cast as computer actionable rules that control the execution of procedures. The procedures are created by chaining computer executable functions (micro-services) into workflows that are executed at remote storage locations. Each consensus is then characterized by a set of policies and procedures. Each user community may choose to enforce a different set of policies and procedures. There is no single set of Technical governance policies. However there is a policy-based framework that enables each community to apply their specific policies."[43]

---

[39] Alva Couch, May 29, 2012.
[40] Shahani Weerawarana, personal email, May 30, 2012.
[41] Reagan Moore, May 25, 2012.
[42] Ibid.
[43] Reagan Moore, May 25, 2012.

In CUAHSI, “Decision [-making] structures are organized around general policies, software policies, and user priorities, respectively. This works well.”[44]

Education and Outreach: Education and outreach should be important part of EarthCube, and particular focus should be on the scientific community. There should be a virtual focus group on the involvement of the scientific community and how to do better outreach.[45]

Evolution of organization: Most organizations start simply. Use cases are beneficial as a means to define scope and simplify the governance process.[46]

Funding: Recommendations include examining the number of paid staff vs. volunteer labor,[47] and reviewing funding mechanisms of various organizations to study best practices and lessons learned.[48]

Goals: EarthCube goals need to be holistic and be focused on much more than just IT issues. A focus on the Grand Challenges research should answer the question “What are we trying to govern?”[49]

Integration and Engagement: Effective governance has an integrative function, and cyberinfrastructure must be a means to better research and science, rather than an end itself. Past NSF activities have demonstrated that someone or something must fulfill a systems integration role to make the parts fit together. Thus, cyberinfrastructure needs to be built with an engineering framework. An example of a governance framework with an integrative function is project GENI [Global Environment for Network Innovations].[50]

Governance issues are associated with large and small science,[51] thus an important question is: How do we integrate with organizations or communities that already have governance structures? (like NASA, for example).[52] A recommendation is to research the strategies that have lasted the longest and have obtained the most results, i.e. the strategies that have made data the easiest to access, use, and maintain.[53]

Interoperability: Model interoperability and data interoperability are two different things.[54] Both present huge challenges in a complex system like EarthCube, and use cases would be beneficial to understanding these issues.

---

[44] Alva Couch, May 29, 2012.
[45] Kerstin Lehnert, community discussion comment, “EarthCube Governance Virtual Plenary Session,” April 11, 2012.
[46] William Michener, April 17, 2012.
[47] Andrew Maffei, community discussion comment, “EarthCube Governance Virtual Plenary Session,” April 17, 2012.
[48] Anonymous, recommendation from Focus Group Webinar, “Engaging Government Agency Partners on Governance,” May 31, 2012
[49] Carroll Hood, community discussion comment, Focus Group Webinar, “Engaging Industry Partners on Governance,” May 31, 2012.
[50] Carroll Hood, May 31, 2012.
[51] Robert Chen, community discussion comment, “EarthCube Governance Virtual Plenary Session,” April 17, 2012.
[52] Ibid.
[53] Anonymous, May 31, 2012.
[54] Cecelia DeLuca, community discussion comment, “EarthCube Governance Virtual Plenary Session,” April 11, 2012.

A recommendation is to create an index for the groups that are trying to promote interoperability, with a focus on design-infrastructure, catalogs, trackers, schemas, etc. [55]

"EarthCube needs mechanisms for technology to interoperate (virtualization mechanisms)."[56]

"The target for governance needs to be on interoperability. What strategies promote interoperability? What standards will enable improved interoperability? What extensibility mechanisms are needed to enable interoperability at the process level, the metadata level, the data level, and the user level? One way to characterize interoperability is through identification of name spaces for objects that are being shared, the identification of the operations that are performed upon each name space, and the virtualization mechanisms that are needed to ensure the operations can be performed across existing community resources. Within the DFC, the name spaces that are managed include users, data, collections, storage resources, state information (metadata), policies, and procedures. The document "Introduction to Layered Architecture" lists the associated operations and virtualization mechanisms."[57]

IT issues: Do vocabularies and ontologies fall into 'soft' ware, or are they separate?[58]

A focus group on software should study governance requirements for software tools.[59]

Leadership: There needs to be defined leadership in EarthCube. The concern that group thinking will yield "impossibly jumbled" criteria is greater than making sure every single person is included in the decision-making process. There needs to be a way to define focus in terms of goals and leadership, and the organization should be strong enough to say no. [60]

Legal Aspects: "There are legal overtones to many governance activities. Consider something as simple as an EarthCube participant collaborating via a mailing list. What are the IP rights that are applicable on each communication? What are the legal strictures that apply to participants in community groups and project teams? What are the legalities present in policies and procedures? How can the governance framework ensure that international trade laws or global IP rights are not violated in the process of collaboration among global EarthCube members?"[61]

"So a governance roadmap should highlight the importance of providing legal guidelines. Most EarthCube participants would be completely unaware of legal aspects and it is up to the Governance framework to protect them, their community, and EarthCube as an entity."[62]

Levels of Governance: There are multiple levels of governance.[63] Governance is necessary in many areas, but it is also necessary at different levels, that is, at the community level, project level, and

---

[55] Ibid.
[56] Reagan Moore, May 25, 2012.
[57] Ibid.
[58] Andrew Maffei, April 17, 2012.
[59] Ilya Zaslavksy, community discussion comment, "EarthCube Governance Virtual Plenary Session," April 11, 2012.
[60] David Fulker, community discussion comment, "EarthCube Governance Virtual Plenary Session," April 11, 2012.
[61] Shahani Weerawarana, personal email, June 1, 2012
[62] Shahani Weerawarana, personal email, June 1, 2012
[63] Sayeed Choudhury, April 17, 2012; Reagan Moore , May 25, 2012, and Shahani Weerawarana, community discussion comment, Focus Group Webinar, "Engaging International Partners on Governance (Pacific Rim)," May 30, 2012.

team level, etc.[64] Thus "Each governance area requires specialized mechanisms/infrastructure/policies etc."[65]

"There are multiple levels of governance. The entire discipline may select the next 10-year research challenges. National scale projects may implement the required infrastructure to collect data, implement solution methods, and make the capabilities available to the discipline. Researchers within the "long-tail" may compare local information with the national initiative results and refine, explore, develop extended result sets. Students may participate in research initiatives under appropriate policy controls. The governance for each of these communities is different."[66]

Metrics: "Governance metrics that are not quantifiable are not needed. For Technical governance, it should be possible to verify each desired property of a collection. For each governance policy, there should be an associated validation policy."[67]

Reproducible Science: "Governance should ensure reproducible science. At a minimum, the information and knowledge required to reproduce a research result has to be provided for each research project. Governance is the process to make this feasible through provision of the necessary infrastructure. If a project is funded, infrastructure must either be provided or developed to enable others to check the result."[68]

"Each research community makes assertions about their research results. A research community should be able to verify each assertion. Governance needs to ensure each community has the infrastructure to validate their own assertions."[69]

Risk Management: [70] "All communication/collaboration plans that are specified in the EarthCube Governance Framework should include explicitly defined "escalation paths." "Many communication and collaboration plans are very detailed regarding regular activities that should happen. However, many forget to specify the escalation path when *things go wrong!* How can any member of the community create an "alert" to enable timely attention on a brewing risk? Who/What is responsible for following up on each alert. What is the escalation path that *leads out of or beyond the immediate communication/collaboration group boundaries*?" [71] "This would also be part of the overall risk mitigation plan in a governance framework."[72]

Scale: Governance scales by two things: 1) Money; and 2) Size of organization and size of partners.[73] "There are multiple national scale data cyberinfrastructure projects, each with a different set of challenges. For instance, the OOI manages real-time sensor data. The National Climatic Data Center [NCDC] manages climate data records. The two systems need to interoperate

[64] Sayeed Choudhury, April 17, 2012, and Shahani Weerawarana, community discussion comment, Focus Group Webinar, May 30, 2012.
[65] Shahani Weerawarana, May 30, 2012
[66] Reagan Moore, May 25, 2012.
[67] Ibid.
[68] Ibid.
[69] Ibid.
[70] Shahani Weerawarana, May 30, 2012
[71] Shahani Weerawarana, June 1, 2012.
[72] Ibid.
[73] William Michener, April 17, 2012.

to enable deposition of OOI climate data records into NCDC. The two systems, however, are based on different infrastructure."[74]

Operational/Managerial, Scientific and Technical Governance: Many individuals and liaisons to other EarthCube Working Groups and Concept teams identified differences between operational/managerial, scientific and technical governance. Some of their responses are included below.

Operational/Managerial Governance

- The Governance "Roadmap may need to include the process of identifying governance areas relevant to EarthCube and the specific governance aspects of each area…"There should be clear differentiation of strategic vs. operational governance structures. The operational structures would necessarily have a greater monitoring & management flavor, whereas the strategic structures would be more focused on guidance and decision-making."[75]

- "There are at least two types of governance: Managerial and Technical. Managerial governance defines the policies for interactions between funding agencies and communities of practice that implement standards and technologies."[76]

- …. For Managerial governance we need mechanisms to ensure sustainability of the infrastructure. This includes evolution of the infrastructure to incorporate new technologies, new approaches towards quantification of knowledge, and new approaches for governance." [77]

Scientific and Technical Governance

- "This distinction [scientific vs. technical governance] is important and should be part of EC governance. It is particularly critical since much of EC technology development will be distributed and will involve federated activities. However, I do think that there needs to be a single overarching entity like the Board of Directors or some other committee that will steer the entire EC effort. Our experience in Unidata is consistent with that. We have a Policy Committee for overall policy decisions and governance, a User's Committee for community engagement and working with our scientific users and educators, and technical advisory committees for our software projects."[78]

- "[Scientific vs. Technical governance is] a useful separation of tasks. For example, attention to SOA [Service-Oriented Architecture], software engineering practices, and data exchange standards would fall more under technical governance, while deciding what problems to solve would fall under scientific governance."[79]

---

[74] Reagan Moore, May 25, 2012.
[75] Shahani Weerawarana, May 30, 2012.
[76] Ibid.
[77] Reagan Moore, May 25, 2012.
[78] Mohan Ramamurthy, personal email, May 22, 2012.
[79] David Arctur, personal email, May 22, 2012.

- “The x-domain interoperability group envisions two kinds of governance: 1) Scientific governance, which determines scientific needs and priorities, [and] 2) Technical governance, which determines shared standards and technologies. Scientific governance includes decisions as to what services will be made generally available, based upon *scientific* merit. Technical governance determines the most convenient technical ways to accomplish that.”[80]

- “I think it is important to separate technical issues from priorities, as described above. Scientists should set priorities, and tech leads should define and manage technologies and their lifecycles.”[81]

- “In the case of X-domain interoperability, there is a natural workflow that starts with compelling use cases (strategic, special-purpose uses) that are generalized into reusable high-level services (general-purpose resources. The primary role of scientific governance is to choose which services should be made generally available, based upon the *scientific* evidence for or against the use cases. Meanwhile, technical governance concerns how standards should be formed and how they should evolve. This is a matter of version control, lifecycle management, and technical decision-making.”[82]

- “Information/data governance would require attention on many aspects including:
    - Information ownership model along with associated ownership roles & responsibilities
    - Definition and identification of information security levels and associated access & usage policies.
    - Information provenance guidelines and enabling infrastructure mechanisms.”[83]

- IT/Cyberinfrastructure governance would require attention on an almost completely different set of aspects [from scientific governance].[84]

- Technical governance defines the policies and procedures that implement a community consensus on assertions about shared collections and processes. The DFC [DataNet Federation Consortium] has chosen to automate Technical governance...”[85]

- “For Technical governance we need mechanisms to define a community consensus on desired properties, policies, and procedures.”[86]

- “The DDMA [Data Discovery, Mining and Access] group has also identified governance (but in the DDMA context) as an activity unto itself. As you can imagine, there are many governance issues at the data level that we feel we need to catch right at the functional level some, if not all, of which might bubble up to the larger governance structure for EC. Even at

---

80 Alva Couch, May 29, 2012.
81 Ibid.
82 Ibid.
83 Ibid.
84 Ibid.
85 Reagan Moore, May 25, 2012.
86 Ibid.

the DDMA level, I can see science vs. technology governance issues (e.g. what level of products should be generated vs. how does one decide where a given data technology is on the hype cycle, etc.)."[87]

Standards: "Standards may promote interoperability for a specific operation on a specific name space. However, new operations continue to appear, new data types are created, new features within data sets are recognized and become relevant, etc. Standards focus on running code, not on new ideas and new research challenges. The governance model needs to allow new standards and new interoperability mechanisms within the infrastructure."[88]

Temporal aspects: "Every aspect of governance should include a temporal component inclusive of associated life-cycles (the process of originating, maturing and retiring particular governance components)."[89]

"Consider governance of policies and procedures. There should be a clearly defined timeline for the definition, operation and retirement of such artifacts. Yet beyond them, practically any component being governed has an associated life-cycle. As an example, consider a project team or a committee."[90]

"Similar to the Apache Software Foundation (http://www.apache.org/), there is a life-cycle process of incubation, graduation and retirement. By associating a pre-specified timeline with such life-cycles, you can ensure timely reviews to ascertain if a component is ready to move on to another stage or not. In the case of named/numbered artifacts, it may result in version changes."[91]

"Thus in general, when putting together a governance framework, it would be useful to require "thinking about" the life-cycle of each aspect/component/artifact and the associated timeline. This will help ensure timeliness, relevance and "liveness" of a governance framework."[92]

Vision: We need further discussion on who determines overall EarthCube goals.[93] The EarthCube vision is not necessarily dominated only by domain scientists, thus it is not helpful to set aside domain scientists as the group who will dominate the EarthCube development vision. The EarthCube vision should be a combination of science and awareness of what technologies are available to implement the vision. Scientists can act as researchers and educators, exploring "The art of the possible."[94]

2.4 FUTURE COMMUNICATIONS

To move EarthCube governance forward, we recommend building upon the community engagement and research review begun as a cornerstone of the Governance Roadmap process. We

[87] Chaitanya Baru, personal email, May 22, 2012.
[88] Reagan Moore, May 25, 2012.
[89] Shahani Weerawarana, May 30, 2012.
[90] Shahani Weerawarana, June 1, 2012
[91] Ibid.
[92] Ibid.
[93] Ibid.
[94] David Fulker, April 11, 2012.

expect community engagement to occur in four steps (for a full description and graphic showing the iterative progression of community engagement as EarthCube grows, see Roadmap Documentation Section 4: Requirements):

1. Identify cyberinfrastructure components of EarthCube
2. Identify the cyberinfrastructure components' organizational paradigms and governance needs.
3. Identify the interaction among and between cyberinfrastructure components and systems within EarthCube.
4. Identify the interactions between cyberinfrastructure components within and outside of EarthCube, and the needs of EarthCube consumers (including those comprising the "long tail" of science)

Community engagement will be coordinated by a community engagement team within the Governance Working Group, whose main task is to engage the community and gather community input to aid in the selection of a governance framework for EarthCube. Once an initial recommendation is made, the proposal will be vetted by the community to determine whether there is sufficient community empowerment to proceed forward with the initial recommendation.

A community engagement program to gather EarthCube governance requirements and vet initial governance framework recommendations may consist of:

1. Gathering requirements for the earthcube.org website
2. Engaging the geosciences, atmosphere, and ocean communities
3. Engaging the computer science, IT and software communities
4. Mapping the EarthCube stakeholder community
5. Developing materials to market EarthCube to potential new users
6. Holding webinars for target communities of practice
7. Engaging existing EarthCube Working Groups and Concept Teams
8. Publishing EarthCube articles in trade journals
9. Engaging domain and IT communities through a stakeholder survey
10. Engaging the community through social media
11. Attending conferences of opportunity, such as ESIP and AGU, to engage potential EarthCube stakeholders
12. Employing public listservs, wikis, and the EarthCube Ning site to communicate with EarthCube stakeholders and provide an additional forum for community input

3.0 CHALLENGES

3.1 GENERAL EARTHCUBE CHALLENGES

Developing a viable organizational and governance structure for any organization can be a challenge. Creating one for multi-disciplined, distributed, virtual collection of scientists, investigators, technologists, system operators, entrepreneurs, and administrators can be nearly impossible unless great care is taken to ensure that the proposed solution is flexible and responsive to meet participant's needs and institutional goals (Figure 2).

FIGURE 2: PREVIOUS WORKSHOPS CONCLUDED THAT CHALLENGES FOR EARTHCUBE ARE MOSTLY CULTURAL AND ORGANIZATIONAL

A plethora of key challenges arose while addressing the EarthCube governance roadmap. Final decisions or recommendations cannot be made by this working group within the given time frame: they must be community driven which by necessity is a lengthier process.

The challenges we consider in this roadmap are not just to creating the governance roadmap per se but also to the role and impacts of a governance process and system on the overall viability and success of EarthCube as a community system. We have framed the challenges so that the community can further discuss and determine viable solutions to these questions.

Key Challenges

1. What shall the enumerated powers/responsibilities of the EarthCube Governance structure be?
2. How will stakeholders interact with the governance structure?
3. How can we avoid short-term thinking in order to meet the long-term EarthCube goals?
4. What are the barriers to participation?
    a. How can governance address the issues brought forth in this roadmap without being perceived as "getting in the way"?
5. How do we build community and engage the end-user?
    a. What does it mean to be a community member? Access to data or involvement in decision making processes or both?

    b. How can community ownership of EarthCube decisions and processes be fostered?
6. How do we best integrate domain science needs with cyberinfrastructure needs?
7. How will decisions be made and what types of decisions must be made?
    a. How will the community participate in these decisions?
    b. What feedback mechanisms are available?
8. How will the governance structure be evaluated?
    a. How can the EarthCube governance framework be evaluated against progress towards EarthCube goals?
    b. What type of metrics or other processes are best suited for this type of evaluation?
9. How will the governance structure evolve?
    a. How can EarthCube enable its own evolution?
    b. How do we recognize and respond to changing community needs?
    c. How do we recognize and respond to technological advances?
10. Is one governance model sufficient to fulfill the governance requirements?
    a. If so, what model is best?
    b. Are multiple models, or "governance as a system" better suited to handle the complex social, cultural, technical, and organizational governance issues that will inevitably arise in an entity as complex as EarthCube?
    c. How can this be determined?
11. What can historical infrastructure dynamics and tensions teach us for the future?
    a. How can governance arrangements address lessons learned from past infrastructure development (see Appendix 1: Research Review, Section 2)?
12. What incentives are there for participation – both from the user community and the content providers?
    a. How can data contributors and cyberinfrastructure builders be incentivized to contribute data/time/resources, etc. to EarthCube?
    b. What processes are available to bring in new contributors?
13. How can the governance framework address the numerous social, cultural, organizational, and technical challenges that accompany the integration of existing systems into the larger EarthCube community?
14. Who leads?
    a. Who is in charge of decision making?
    b. What decisions need to be made?
    c. Does that individual/group in charge of making decisions change depending on what type of decision is being made?
    d. Who has the ability to resolve disputes/willingness to say no?
    e. How can we avoid group think?
    f. What governance mechanisms can determine leadership roles and the processes by which those roles may be filled?
    g. What mechanisms are in place to allow for new leadership to emerge if/when necessary?
15. What are the objectives and goals of the governance framework?
    a. Who is responsible for creating and maintaining these objectives?
    b. What governance mechanisms can be built into the governance framework to be aligned to meet these objectives?
    c. How can we create mechanisms to evaluate the EarthCube governance framework against progress towards EarthCube goals?

16. How do you determine responsibility and accountability?
    a. How can the governance framework help determine who is responsible for what tasks?
    b. How can governance foster a sense of responsibility and accountability among EarthCube participants?
17. How will the governance framework be sustained?
    a. How can EarthCube be sustained financially?
    b. How can EarthCube be sustained in terms of community support and end-user engagement?
18. How do you build trust?
    a. How can a sense of trust be fostered among community members in general?
    b. How can data contributors trust that their data will be handled appropriately, and that they receive proper credit for their contributions?
    c. How can data users trust the validity and plausibility of data contributed to EarthCube?
    d. How do you ensure that EarthCube does not compromise the primary missions of the domain-specific groups by siphoning resources away from serving the needs of domain users?
19. How do you manage risk?
    a. How do you deal with unforeseen or unwanted outcomes?
20. How do you manage legal challenges?
    a. How can intellectual property rights be protected?
    b. How can proper data use, citation and publication be promoted?
    c. What happens if a group refuses to go along with certain system/network requirements, even if adopted by community consensus or imposed by system authorities?
21. Who determines the EarthCube vision?
22. How can scientific needs be addressed by available technologies?
23. How can allegiance to primary domains be mitigated?
    a. Domain groups may owe primary allegiance to specific domains and EarthCube could compromise the primary missions of the domain-specific groups by siphoning resources away from serving the needs of domain users.

3.2 CONCEPTUAL AND PROCEDURAL CHALLENGES

The evolution of all infrastructures we examined show that they have common characteristics of starting out with interplay among largely independent and even disparate elements, that grow, link with other elements, and create and manage interactions among the components. Such is the current path for cyberinfrastructure.

The history of infrastructure development (rail, telephone, highway system, electrical grid, etc.) has destroyed any notion that it is a planned, orderly, and mechanical procedure.[95] As Edwards et al. discovered, "Robust cyberinfrastructure will develop only when social, organizational, and cultural

[95] Edwards et al., "Understanding Infrastructure," i.

issues are resolved in tandem with the creation of technology-based services."[96] Thus, there is no single or correct path to link the social/cultural and technological aspects of building a cyberinfrastructure.

Edwards et al. identified four common infrastructure dynamics, and eight common tensions, all of which can applied to EarthCube. Common infrastructure dynamics include:[97]

1. **Reverse Salients:** critical, unsolved problems
2. **Gateways.** "Plugs and sockets that allow new networks to be joined with existing networks with minimal constraint."
3. **Path Dependence.** Early choices can constrain available options later on in the development process.
4. **Scale Effects:** It takes 40-50 years for short-term, local infrastructure projects to grow into large-scale, functional infrastructures.

Edwards et al. and his colleagues summarized infrastructure tensions in terms of:[98]

1. **Time, Scale, and Agency.** Time (short-term funding decisions versus the long-term time-scale needed for infrastructures to grow); Scale (choices between worldwide interoperability and local optimization); Agency (how to navigate planned versus emergent change).
2. **Winners and Losers.** Those who benefit from, and those who are left behind, as a result of certain decisions or the course of infrastructure development.
3. Interest and Exclusion. The unequal distribution of power and consequences of infrastructural change among stakeholders and contributors.
4. **Ownership and Investment.** Tensions of ownership and investment occur when changing infrastructures run into intellectual property rights, public/private investment models, and policy objectives.
5. **Data Cultures and Data Tensions.** Data storage, preservation, and curation are at the center of some of the most complex challenges to CI development.
6. **Data Storage and Meaning.** The storage, preservation and curation of data.
7. **Data Sharing.** Data sharing necessitates tools that have the ability to accommodate and translate different needs from distinct domain communities.
8. **Trust.** Data trust is a key component of data sharing and involves data generators and data users.

Issues of ownership, management, control, and access are pervasive among all the Working Group and Concept Team roadmaps for EarthCube, and examples previously compiled are being repeated in the current EarthCube scoping process. For example:

- Who decides on rules and conventions for sharing, storing, and preserving data?
- Local variation vs. global standards: How do we resolve frictions between localized routines and cultures that stand in the way of effective collaboration?

How can national cyberinfrastructure development move forward without compromising possibilities for international or even global infrastructure formation?

[96] Ibid.
[97] Ibid., 14-19
[98] Ibid., 24-33.

3.3 SOCIAL AND CULTURAL CHALLENGES

Motivations and Incentives

The EarthCube Cross-Domain Interoperability Roadmap Concept Team identified several social and cultural challenges. A strong cultural paradigm is that scientists have traditionally had a strong motivation not to share data. This is primarily in order to protect one's rights to publish first, particularly after expending valuable time and effort to create a given data source. Key questions include ownership of data acquired using public funding, and under what conditions it is ethical for such data to be withheld. They proposed a solution to this problem is to create incentives for data creators to document and share their data, or at least the metadata describing the data.

Self-Selected or Closely Held Leadership

> *One database to rule them all,*
> *One database to find them,*
> *One database to bring them all,*
> *And in the darkness bind them.*
> *With apologies to J.R.R. Tolkien*

FIGURE 3: FOR A DECADE, THE COMMUNITY HAS ASKED WHO GOVERNS CYBERINFRASTRUCTURE AND HOW MUCH CONTROL WILL THEY HAVE.

During the past decade, there have been a number of geoinformatics workshops that heard repeatedly from the community that vesting power in any form of "data czar" was not acceptable. Concerns have been about perceived lack of community involvement or control, and imposition of standards or requirements that could necessitate significant expense to adopt or implement and would not be optimal for all parties, among others (Figure 3).

Additional challenges raised by the Apache Software Organization concern any community which centers on a few individuals who may work virtually uncontested. The Foundation determined this to be detrimental to quality, stability, and robustness of both of their core functions of code and long term social structures.

A related challenge is that of relying on volunteers, even for "owners" of services or processes in the system. Do volunteers represent the community adequately, rather than primarily their own interests? Relying on self-selected volunteers is inherently a risk: can cyberinfrastructure even get the system built, let alone having it achieve what we want it to do?

Levels of Participation

The current involvement in EarthCube is from individuals, groups, organizations, institutions, networks, and systems, all as contributors and "owners" of EarthCube and as consumers of services and data supplied via EarthCube. This is not expected to simplify over time. Each one of the stakeholder groups (contributors, participants, and users) will have their own roles and needs that have to be considered in designing the governance implementation process and governance system(s).

Another challenge is the sheer number of Working Groups and Concept Teams holding concurrent and often overlapping webinars, calls, and work sessions, as all groups rushed to complete their own roadmaps by the deadlines. This limited participation by many who might have otherwise been centrally engaged in corollary activities.

Despite these efforts, the number of people engaged in EarthCube is still only a miniscule fraction of the overall geoscience and IT communities. However, it can be argued that penetration is significant among the most active cyberinfrastructure proponents. One of the challenges of completing this roadmap was the unprecedented short time frame approximately 3 months from concept initiation to completion and delivery. This precluded use of many of the traditional and tested communication methods to engage larger communities, such as posting announcements in professional newsletters, newspapers, and magazines and circulating announcements through professional networks. The reliance on online communication techniques may have skewed the population demographics contacted as well as omitting those with greater degrees of separation from the core EarthCube constituencies.

Going forward, successors to Governance Working Group and the other EarthCube groups may be able to learn from our collective experiences of writing the roadmap and engaging the community, and may have additional time to reach out further.

Types of organizations

In addition to the levels of participants, attention must be given to the types of organizations that are and will be part of EarthCube, including academia, government, industry, NGO's, international groups and bodies, and professional associations. Each entity brings a set of opportunities and challenges that should not be overlooked. A challenge is ensuring that the roadmap does not implicitly discourage or prevent organizations from participating or contributing in the optimum manner.

Collaboration among Domain Scientists and IT Specialists

Beyond the challenges of collaboration of domain scientists, technologists, and social scientists, there are challenges of engaging and including other stakeholders, particularly those outside of the largely self-selected EarthCube community, and creating a governance system that accommodates participation ranging from individuals to groups or teams, to organizations, institutions, and networks. Tensions exist also between what is needed and what can be achieved within existing structures, budgets, and intellectual resources.

3.4 TECHNICAL CHALLENGES

We expect to identify challenges from all segments of EarthCube. For example, the EarthCube Cross-Domain Interoperability Concept Team concluded "there is no technological challenge to achieving this degree of interoperability, simply the resources and community engagement and awareness to assemble and to maintain the catalogs, to implement standards, and to govern the operation of such services." This conclusion seems to exist across the breadth of EarthCube components.

Examples from Use Cases
Additional challenges have come to light via use cases.

User case: Centralized versus distributed data storage
In large complex initiatives, especially where there are multiple domains potentially making competing demands, there will be conflicts and tensions between community-driven governance models and a more central master structure to allow for a quantitative uniform decision-making process. Some will argue for a type of framework, similar to the LSST (Large Synoptic Survey Telescope) approach, where a committee of some type will be used to construct a set of "sizing guides" that capture, within the infrastructure design, the actual dollar cost of any individual hardware or software component.

The plans for a centralized governance structure provide concrete criteria through which to evaluate every component of the system holistically to determine the order of development and deployment efforts, manage risk, and ensure that integration efforts meet total project objectives and schedule. Such formal project management documentation, while seemingly burdensome compared with the more fluid process typically used in single-investigator projects, is absolutely critical when coordinating a large distributed group consisting of diverse expertise and geographically remote teams and individuals. A cost analysis needs to be completed on any development of open-source tools when commercial tools are more mature architecturally and easier to use.

The interpretation and evaluation of the cost analysis is difficult considering that EarthCube would need to evaluate the needs of multiple domains.

There are two significant models for large data storage projects: distributed data (also referred to as the federated model) and centralized data. Both of these data models have significant benefits and issues. The data distributed model is initially the least costly option and can be implemented quickly. In a distributed model the data is physically distributed across multiple sites (nodes) that each specialize on one of the geoscience domains, while a smaller number of centralized nodes contain the metadata for searching. One of the benefits of distributed system model is that sites can each specialize in a single science domain and new resources can be added quickly. However, this ease of addition also makes it more vulnerable to data loss: the world wide web succeeded because of the ease of adding new web servers to the network, but the majority of the content posted to the web since its creation has subsequently been lost.[99] In addition, the movement of large datasets across the network can be problematic.

[99] Kalev Leetaru, "A Decade and a Half of Archiving the Web for Data Mining: Lessons Learned and How Users Use Web Archives," International Internet Preservation Consortium, 2012.

In the short term the distributed data model is probably the least expensive and most rapidly implemented because one can use the current servers and distribution centers that are already in place. In the longer term it could be more a more expensive option because of the need to keep multiple nodes active. A centralized model for data management has been the preferred choice among the largest data driven projects such as LSST and GenBank. However, it is important to note that in both cases the data collection, archiving, retrieval, and analysis is from a single scientific domain. The governance process needs to evaluate how a centralized depository could work with diverse scientific domains. There would be advantages to having a centralized data model because data movement issues would be less of concern. The largest projects, such as LSST, have reached a point where data has become too large to be moved, and so computation on the largest datasets will have to move to where the data is. Based on the LSST model, ultimately a local compute cluster collocated with the data will allow allocated computation directly across the largest datasets.

The EarthCube governance implementation has to be flexible to accommodate either or both data storage models and a system that will be evolving. It is anticipated that it would take years to include all of the data from the current distributed nodes in a central database. A set of centralized servers would provide data storage and computing resources, while remote nodes would allow for specialized portals and legacy data. How does the governance structure change if the data model changes from distributed to a more centralized model or vice versa?

Use Case: NEES Grid
One of the lessons reinforced from the NSF NEESGrid project was that a successful effort must involve strong collaboration between both domain scientists and technologists. Scientists need to have a strong role in developing the driving vision but the technologists must also have a strong role in order to shape the vision to be computationally tractable. The dangers of a totally technology-driven approach are seen in the caBIG© program where costly technologies were developed without the ongoing input of the scientific community (in this case cancer researchers), resulting in a collection of tools that ultimately did not find widespread adoption.

3.5 TRENDS AND DRIVERS

The EarthCube governance framework should optimally be developed early in the creation of EarthCube. The governance team must include all members of EarthCube community in deciding the type of conceptual architecture(s) which will be needed to integrate advanced information technologies that facilitate access to resources, such as computational tools and services, instruments, data, and people. As a result, it is the suggestion of the Governance Working Group that the governance framework for EarthCube be in place prior to the implementation of initial technical and scientific components, as the governance structure and processes will dictate how the other aspects of the system function and integrate.

Federal Government Initiatives
External drivers will likely have tremendous impacts on the construction, adoption, and sustainability of cyberinfrastructure. The White House "Big Data" and Digital Government (http://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government.html?goback=.gde_3427896_member_120023144) initiatives are just a couple of

examples. Digital Government includes an explicit endorsement of web APIs and services ("Make Existing High-Value Data and Content Available through Web APIs"), and lays out an aggressive timeline that includes creating a catalog of Web APIs from different agencies and requiring all new IT systems to expose Web APIs - both within 12 months. There is a significant focus on intra-agency governance for digital services, and on standards. There is also a specific provision for a new "Digital Services Information Center" to provide support to help agencies develop Web APIs (within 6 months). There seem to be many parallels to the EarthCube agenda, and perhaps a framework for interacting with government stakeholders.

Similarly, over the last several years the U.S. Department of Defense (DoD) has moved in the direction of a network-centric data sharing paradigm that is very similar to what NSF is proposing for EarthCube/CIF 21. DoD. As part of this effort, DoD attempted to jump start an inter-agency and coalition effort for an integrated Maritime Domain Awareness (MDA) capability, but money ran out and it withered. However, the MDA community of interest for the DOD Global Information Grid (GIG) suggested that the lessons learned could have relevance for us.

Cloud Computing

There is also an emphasis on cloud computing. The EarthCube exercise might do a great service by simply specifying what types of cloud services are needed for cyberinfrastructure and under what business model(s). As one of our correspondents wryly noted, for many practitioners, “you will have to pry my server from my cold, dead hands.”

3.6 REMAINING ISSUES

Attention must be given to the types of organizations that are and will be part of EarthCube, including academia, government, industry, NGO’s, international groups and bodies, and professional associations. Each brings a set of opportunities and challenges that should not be overlooked. A challenge is ensuring that the roadmap does not implicitly discourage or prevent organizations from participating or contributing in the optimum manner.

The Governance Working Group did not come to consensus on whether governance of EarthCube requires a single, central management body, responsible for prioritizing allocation of resources, coordination and communication among all participants, and distributing funds, or if EarthCube governance will be a distributed system, similar to other types of infrastructure, such as the Internet, with a loose to closely coupled network of organizations that work collectively. We expect that this is a fundamental dichotomy that is not simply confined to differences between scientists and technologists, for example, and will need to be discussed further as EarthCube develops.

There are challenges to developing an effective EarthCube roadmap inherent in the current process given the limited time frame. Among these challenges:

1. Comprehensive background research review of governance topics from the domain sciences, IT, and social sciences is not yet complete.
2. We identified many governance models, but have not been able to fully evaluate them.
3. Further work is needed to evaluate the pros and cons of different models and determine which may be suitable for EarthCube.

4. Our knowledge of the other EarthCube Working Group and Concept Team governance issues and needs is not yet complete.
5. We have yet to fully engage the broader Earth, information, and IT science communities, thus our knowledge of their governance needs is limited.
6. There is limited information available about problems and failures of past projects that we can incorporate as things to avoid.

Despite this intimidating list of challenges, there is a growing realization that the initial roadmap is not required to solve all of them. In fact, the June 2012 NSF Charrette provided an ideal venue to address and try to resolve many of them in consultation with the other Working Groups and Concept Teams. Post-Charrette activities to revise the roadmap and develop the Governance Framework document have served as additional venues to begin to address some of these challenges.

4.0 REQUIREMENTS

4.1 COMMUNITY ENGAGEMENT PROCESS

The next step in developing the EarthCube governance framework is to identify the current components of cyberinfrastructure (data and service providers), their organizational paradigms, governance needs, the interactions needed among them, and between them and systems outside of EarthCube, and the needs of EarthCube consumers (including those comprising the 'long tail' of scientists. We propose an iterative outreach and requirements-gathering process that builds upon the initial community engagement plan that informed the writing of the Governance Roadmap. This process is described below.

Steps Taken Thus Far

Step 1: Identify cyberinfrastructure components of EarthCube. NSF's identification of the initial cyberinfrastructure components of EarthCube, divided into the Working Groups (Governance, Ontologies & Semantics, Workflow, and Data Mining, Access and Discovery), and the Concept Awards (Web Services, Cross-Domain Interoperability, Earth System Model, Layered Architecture, and Brokering) formed the first step in identifying the cyberinfrastructure components of EarthCube.

Step 2: Identify the cyberinfrastructure components' organizational paradigms and governance needs. Organizational paradigms and governance needs of current cyberinfrastructure components were identified through a governance research review and solicitation of community feedback from EarthCube Working Groups and Concept Teams. Background research informed the solicitation of feedback from EarthCube Working Groups and Concept Teams, while community input guided the direction of the research review.

Step 3: Identify the interactions among and between cyberinfrastructure components and systems within EarthCube. Input from within the EarthCube community was synthesized and cross-referenced to gain a more complete picture of governance wants and needs, and interactions among and between, current cyberinfrastructure components.

Step 4: Identify the interactions between cyberinfrastructure components within and without EarthCube, and the needs of EarthCube consumers (including those comprising the 'long tail' of scientists). The scope of the research review and solicitation of community feedback expanded in order to engage potential stakeholders outside the initial EarthCube community, including representatives from industry, international organizations, and government agencies.

We are currently in implementing step 4 of the first spiral in the development of the governance framework (Figure 4). The initial cyberinfrastructure components have been identified (Working Groups and Concept Teams – Step 1), and their organizational paradigms and governance wants we analyzed via their draft roadmaps (Step 2). Following the June 2012 Charrette, we analyzed the identified the interactions among the Working Groups and Concept Teams by cross-referencing their Roadmaps (Step 3), and we are in the processing of an aggressive community engagement plant to reach out to long-tail scientists and industry, international and government agency representatives outside of the initial EarthCube community (Step 4).

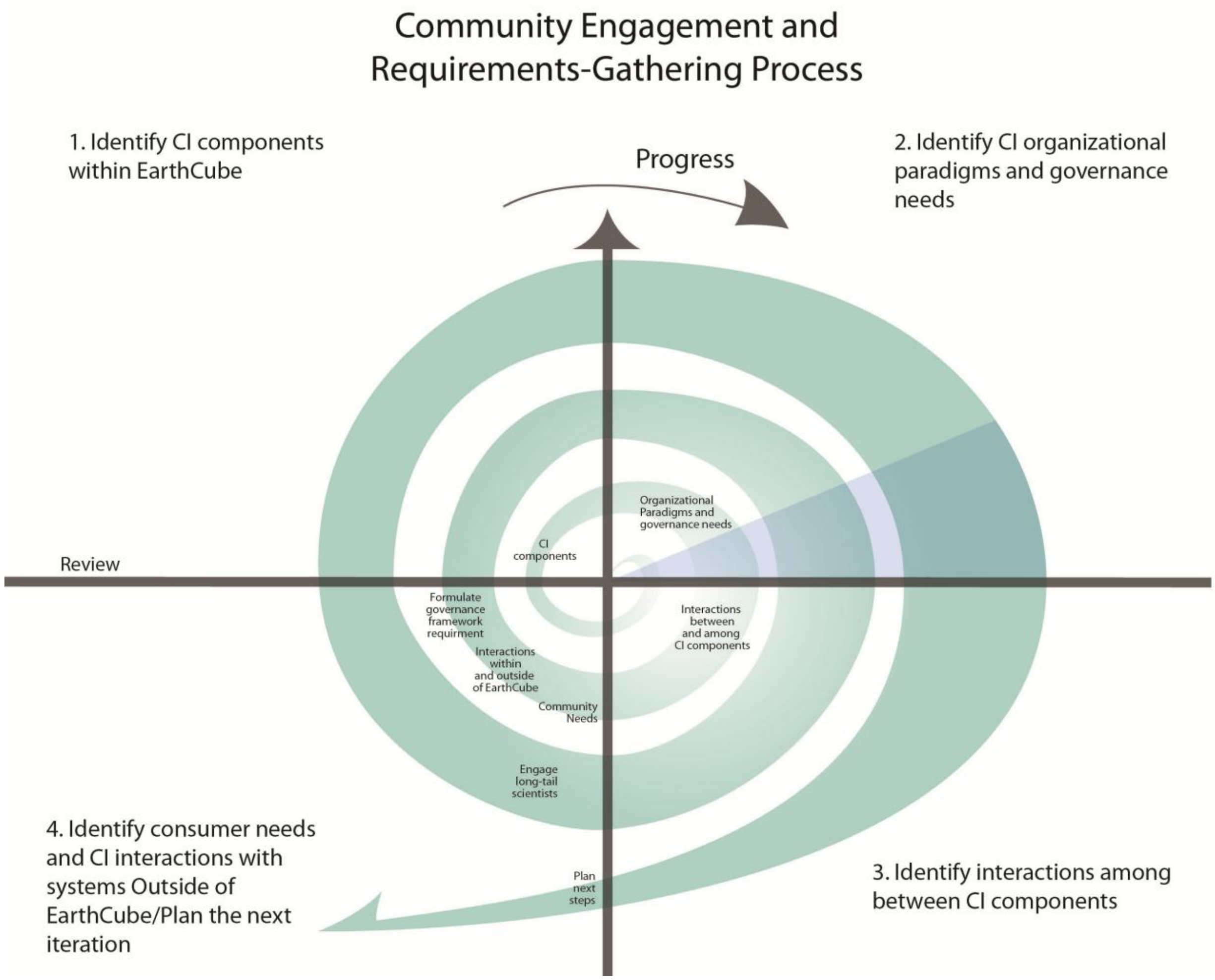

FIGURE 4: ITERATIVE REQUIREMENTS-GATHERING PROCESS FOR EARTHCUBE GOVERNANCE.

Next Steps

The second spiral in the development of the EarthCube governance framework began after the June 2012 Charrette, as additional cyberinfrastructure components, EarthCube stakeholders and their organizational paradigms and needs were identified. This iterative process will continue on an increasingly larger scale as the EarthCube community expands and new cyberinfrastructure components within EarthCube are identified.

4.2 CURRENT EARTHCUBE GOVERNANCE REQUIREMENTS

The requirements presented below were generated by community engagement and requirements-gathering process discussed in Section 4.1. This list will surely grow with the identification of additional cyberinfrastructure components and their respective governance needs.

Best Practices and Lessons Learned: EarthCube governance should take into account the best practices and lessons learned from past cyberinfrastructure, including common dynamics and tensions, as described by Edwards et al. in "Understanding Infrastructure." EarthCube's governance framework should generalize from specific success stories to develop new general-purpose services.

Barriers to Participation: EarthCube governance should be designed so that anyone who wants to participate may do so. The governance framework should recognize and mitigate potential barriers to participation by industry, government agencies, academia, international partners, students, educators, policy makers, long tail scientists and others**.**

Collaboration: EarthCube governance should provide ground rules on collaboration and properties the collaboration environment will maintain.

Communication: EarthCube governance must clearly define the communication mechanisms that will solicit input for decision-making, disseminate results of the decision-making processes, communicate changes in governance, and solicit general feedback from community members. The development of a governance approach for EarthCube will require communication between all stakeholders. Communication mechanisms should promote transparency and communication across the EarthCube community.

Community: EarthCube governance must be community-based, and must meet the short-term and long-term needs of the community. The governance framework must clearly define the EarthCube community, how the community-building process is conducted and when it must begin and what it means to be a community member (in terms of access to data, leadership, involvement in decision-making processes, ability to give feedback, etc.), and what processes are in place to collect and respond to feedback.

Culture: EarthCube governance should take into account the cultural and organizational norms of the different groups and communities that will be part of the overall EarthCube community.

Decision-making processes: EarthCube governance should determine what types of decisions must be made, and clearly define the decision-making processes correlating to each type of decision. Governance mechanisms should be put in place to break decision-deadlocks/have the authority to resolve disputes, taking into account disparate stakeholder requirements, resource considerations and other constraints, to define the optimal path forward for EarthCube.

Design: EarthCube governance should complement EarthCube's design and vice-versa. The governance model(s) and mechanisms employed in the EarthCube governance framework should be appropriate for EarthCube design and technological structure.

Education and Outreach: EarthCube governance should incorporate ongoing education and training, in addition to general community engagement. The full potential of EarthCube capabilities will not be realized if community members are not aware of, or do not understand, tools and resources available. Conversely, a lack of education and outreach could lead to EarthCube failure, or at the very least, reduce EarthCube's effectives due to lack of training and education of the community for whom EarthCube will be created.

Evaluation: EarthCube governance must build-in mechanisms to periodically or continually evaluate the governance framework against overall progress towards achieving long-term goals. There must be clear indicators of success against which the governance framework can be analyzed.

Evolution: EarthCube governance should be dynamic and enable its own evolution. Governance mechanisms that recognize and respond to changing community needs and take advantage of emergent technologies should be built into the EarthCube governance framework.

Goals and Objectives EarthCube governance must clearly define overall EarthCube goals and objectives, including who is responsible for creating and maintaining these goals, and what processes may be employed to update these goals as necessary.

Governance as a System and Multiple Levels of Governance: EarthCube governance should recognize that governance is not a one-size-fits model. Different governance archetypes and different levels of governance will likely be needed to govern different aspects of EarthCube. For example, several governance components that should be taken into account include operational/managerial governance, enterprise-level vs. internal governance (governance within each CI component, community of practice, etc.), scientific vs. technical governance, etc.

Integration and Interoperability: EarthCube governance should address the technical, cultural and organizational issues related to integration and interoperability of data, technology, and existing communities of practice (domain scientists and IT experts, government agencies, industry, academia, scientific consortiums, students, educators, policy makers, and other potential stakeholders) into the EarthCube network and community. Governance mechanisms should focus on social, organizational, cultural, and technical solutions, or some combination of all of these, to deal with issues of integration. An important question is: how does EarthCube integrate with existing communities of practice that already have governance structures?

Leadership: EarthCube governance must specify leadership roles, and create mechanisms to determine who can fill these roles, who can make decisions, and who has the authority to resolve disputes. EarthCube's governance framework should employ mechanisms that ensure diverse representation of the EarthCube community in decision-making and strategic priorities/goal creation processes. The governance framework should allow for new leadership to emerge if/when necessary.

Legal Issues: EarthCube governance should consider the legal components of building an EarthCube community and should protect individuals, their communities, and EarthCube as an entity.

Metrics: EarthCube governance should include quantitative and qualitative metrics to evaluate progress towards overall EarthCube goals.

Reproducible Science: EarthCube governance should ensure reproducible science and the ability of researchers and research communities to verify the validity of their results.

Risk Mitigation: EarthCube governance should examine the possible risks associated with EarthCube and create plans for when things go wrong. Mechanisms should allow community members to bring a growing risk to attention so that it may be addressed in a timely fashion. Mitigation plans should be in place to identify who or what is responsible for following up on risks.

Scale: EarthCube governance must be scalable, both in terms of funding and in terms of size and organization, and size of partner networks.

Sustainability: EarthCube governance must assure EarthCube continuation and stability while maintaining flexibility as new technologies, business models, and user needs emerge. Governance should address specific issues of sustainability, including funding mechanisms, community support/buy-in, and education/training. The governance framework should build-in mechanisms that encourage community participation and contribution, including means to make EarthCube perceived as valuable, interesting, useful, and beneficial to existing community members and potential contributors. EarthCube governance must address issues of long-term funding sources.

Temporal Aspects: EarthCube governance should consider the lifecycle of each aspect/component/artifact and the associated timeline.

Trust: EarthCube governance should foster a sense of trust among data and service providers, scientists, and technologists. Governance mechanisms should establish procedures ensuring data and service providers receive proper credit for their contributions, and that appropriate citation and publication procedures are followed. Data-and software users should be able to trust the quality of data and services available; thus the governance framework should create a shared notion of quality assurance to evaluate the validity of different sources.

Vision: EarthCube governance should take into account scientific priorities and existing technologies (and social constructs) to craft and implement the EarthCube vision.

These requirements will be expanded upon and categorized in the Governance Framework to be presented to NSF and the EarthCube community by the end of August 2012 (Step 1 of the Milestones and Tasks).

5.0 STATUS

5.1 GOVERNANCE DEFINITIONS

Weill and Ross define IT governance as "Specifying the decision rights and accountability framework to encourage desirable behavior in using IT. IT governance is not about making specific IT decisions–management does that–but rather determines who systematically makes and contributes to those decisions."[100]

The Organization for Economic Cooperation and Development defines governance as "Providing the structure for determining organizational objectives and monitoring performance to ensure that objectives are attained."[101]

The EarthSystem Commodity Governance Project defines governance as "align[ing] an organization's practices and procedures with its goals, purposes, and values. Definitions vary, but in general governance involves overseeing, steering, and articulating organizational norms and processes (as opposed to managerial activities such as detailed planning and allocation of effort). Styles of governance range from authoritarian to communalist to anarchical, each with advantages and drawbacks."[102]

In open source software projects, the "governance model describes the roles that project participants can take on and the process for decision making within the project. In addition, it describes the ground rules for participation in the project and the processes for communicating and sharing within the project team and community."[103]

5.2 IT GOVERNANCE

Effective governance is defined as an "Actively designed…set of IT governance mechanisms (committees, budgeting processes, approvals, and so on) that encourage behavior consistent with the organization's mission, strategy, values, norms, and culture."[104] Effective governance gives the organization the ability to recognize and respond to changes, and encompasses three questions:[105]

1. ***What decisions must be made to ensure effective management and use of IT?***
2. ***Who should make these decisions?***
3. ***How will these decisions be made and monitored?***

---

[100] Weill and Ross, *IT Governance*, 2
[101] Organization for Economic Cooperation and Development, *OECD Principles of Corporate Governance*, (April 1999), 5, 219, quoted in Weill and Ross, *IT Governance*, 4-.5
[102] "Governance," EarthSystem Commodity Governance Project, last modified 2012, http://earthsystemcog.org/projects/cog/governance_object.
[103] Ross Gardler and Gabriel Hanganu, "Governance Models."
[104] Weill and Ross, *IT Governance*, 2-3.
[105] Ibid, 10.

Question 1: What decisions must be made? IT governance decisions include:[106]

1. **IT principles**: Clarifying the business role of IT
2. **IT architecture**: Defining integration and standardization
3. **IT infrastructure**: Determining shared and enabling services
4. **Business application needs**: Specifying business needs for purchased or internally developed IT applications
5. **IT investment and prioritization**: Choosing which initiatives to fund and how much to spend

Specific decisions within each category are described below:[107]

1. **IT Principles**
    a. What is the enterprise's operating model?
    b. What is the role of IT in the business?
    c. What are IT-desirable behaviors?
    d. How will IT be funded?
2. **IT architecture**
    a. What are the core business processes of the enterprise?
    b. How are they related?
    c. What information drives these core processes? How must the data be integrated?
    d. What technical capabilities should be standardized enterprise-wide to support IT efficiencies and facilitate process standardization and integration?
    e. What activities must be standardized enterprise-wide to support data integration?
    f. What technology choices will guide the enterprise's approach to IT initiatives?
3. **IT Infrastructure**
    a. What infrastructure services are most critical to achieving the enterprise's strategic objectives?
    b. For each capability cluster, what infrastructure services should be implemented enterprise-wide, and what are the service-level requirements of those services?
    c. How should infrastructure services be priced?
    d. What is the plan for keeping underlying technologies up to date?
    e. What infrastructure services should be outsourced?
4. **Business Application Needs**
    a. What are the market and business process opportunities for new business applications?
    b. How are experiments designed to assess whether they are successful?
    c. How can business needs be addressed within architectural standards? When does a business need justify an exception to a standard?
    d. Who will own the outcomes of each project and institute organizational changes to ensure the value?
5. **IT Investment and prioritization**
    a. What process changes or enhancements are strategically most important to the enterprise?
    b. What are the distributions in the current and proposed IT portfolios? Are these portfolios consistent with the enterprise's strategic objectives?

---

[106] Weill and Ross, *IT Governance*,Ibid., 10-11.
[107] IbidWeill and Ross, *IT Governance*, 54-55.

c. What is the relative important of enterprise-wide versus business unit investments? Do actual investment practices reflect their relative importance?

Question 2: Who makes decisions? [108]

1. **Benevolent Dictatorship:** benevolent dictator
2. **Business monarchy**: top managers
3. **IT monarchy**: IT specialists
4. **Feudal**: each unit makes independent decisions
5. **Federal**: combination of corporate center and business units with or without IT people involved
6. **IT duopoly**: IT group and one other group (ex: top management or business unit leaders)
7. **Anarchy**: isolated individual or small group decision making
8. **Meritocracy**: distributed control awarded based on contributions to the project.

Several governance archetypes are available to make IT-related decisions (Table 1).

IT Governance Archetypes

STYLE	WHO HAS DECISION OR INPUT RIGHTS?
Benevolent Dictatorship	Benevolent dictator
Business Monarchy	A group of business executives or individual executives. Includes committees of senior business executives (may include CIO). Excludes IT executives acting independently
IT Monarchy	Individuals or groups of IT executives
Feudal	Business unit leaders, key process owners or their delegates; each 'fiefdom.'
Federal	C-level executives and business groups (e.g., business units or processes); may also include IT executives as additional participants. Equivalent of the central and state governments working together
IT Duopoly	IT executives and one other group/ Interactions between any two system elements
Anarchy	Each individual user

TABLE 1: "IT GOVERNANCE ARCHETYPES"[109]

These governance models are represented as graphics in Figure 5 below.

[108] Weill and Ross, *IT Governance*,Ibid., 12.
[109] Adapted and modified from MIT Sloan School Center for Information Systems Research (CISR), 2003, Reproduced from Weill and Ross *IT Governance*, 2004, 59. Benevolent Dictatorship model from Ross Gardler and Gabriel Hanganu, "Governance Models."

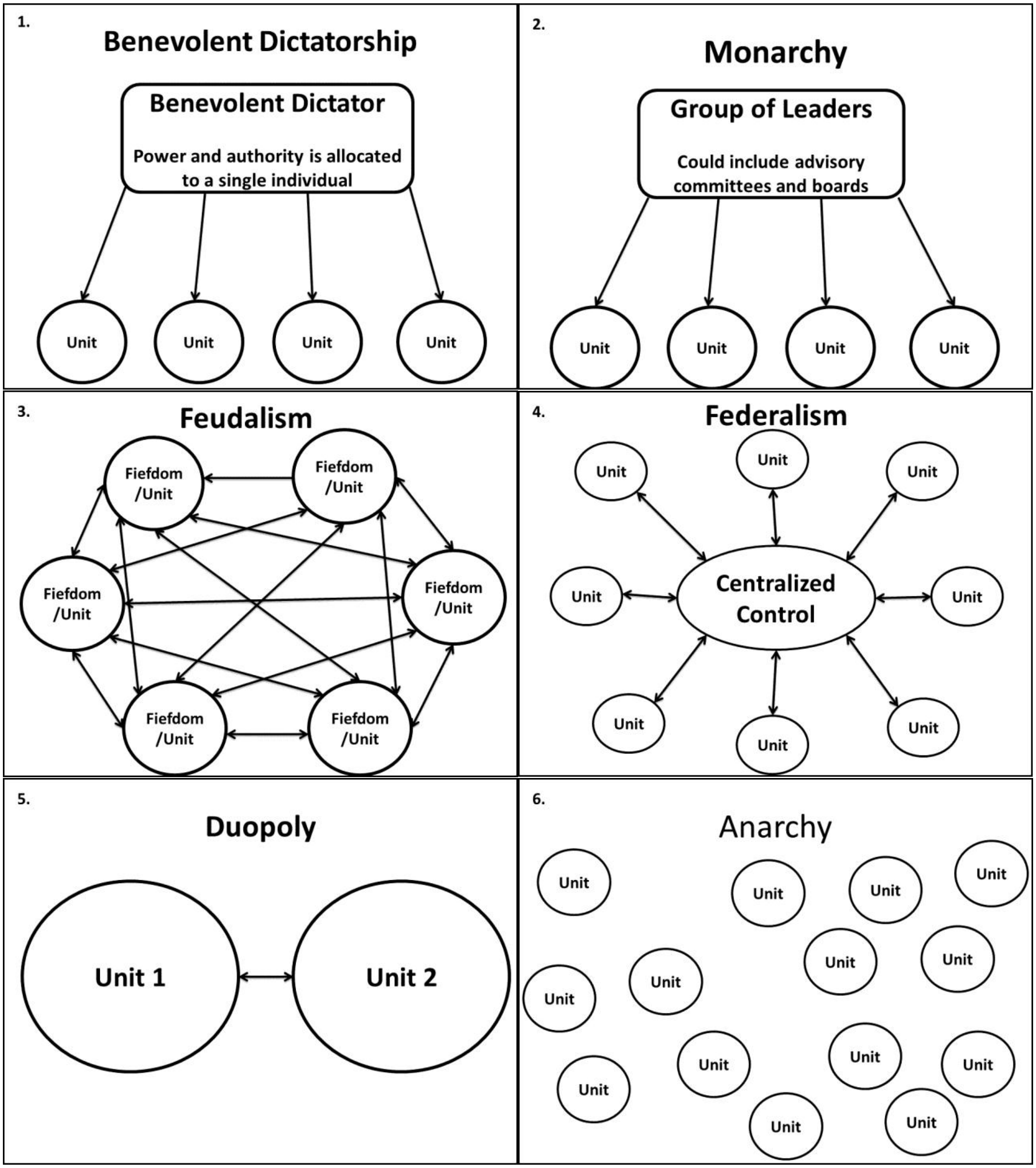

FIGURE 5: GRAPHICAL REPRESENTATIONS OF SEVERAL GOVERNANCE MODELS, ORDERED ON A SCALE OF MOST CENTRALIZED (BENEVOLENT DICTATORSHIP) TO MOST DECENTRALIZED (ANARCHY).[110]

[110] Business and IT Monarchy models are represented simply as Monarchy. Meritocracy can be employed to select groups or individuals to make decisions in any one of these models, therefore it is not represented as a stand-alone governance model.

Question 3: How will decisions made and be monitored? Weill and Ross argue that one governance archetype or model may not appropriate for all decisions. Instead, they view governance as a system of archetypes within an organization, in which decisions are made according to the most appropriate archetype for that type of decision (Table 2). Ideally, each organization's decision-making structures are aligned towards achieving that organization's goals.

Governance as a System Arrangements Matrix

Decision Archetype	***IT Principles***	***IT Architecture***	***IT Infrastructure Strategies***	***Business Application Needs***	***IT Investment***
Business Monarchy					
IT Monarchy					
Feudal					
Federal					
Duopoly					
Anarchy					

TABLE 2: "GOVERNANCE AS A SYSTEM ARRANGEMENTS MATRIX—WHICH GOVERNANCE ARCHETYPES ARE USED FOR DIFFERENT TYPES OF DECISIONS?"[111]

How Enterprises Govern

Weill and Ross found a number of common governance patterns in their research. The found that firms gathered input from many sources, and decision-making rights were allocated to different groups depending on the decision being made. The Federal model was most common for business-oriented decisions, and committees, budgets, and cross-functional process teams to provide input and feedback were common governance mechanisms.

The IT Duopoly was often used for less technical decision-making, focusing on decisions regarding IT principles, business application needs, and IT investment, but technical decisions were commonly made via an IT Duopoly and Federal System. Weill and Ross' findings are summarized in Table 3 below. Each decision is divided into two processes: information gathering and decision-making.

[111] MIT Sloan School Center for Information Systems Research (CISR), 2003, Reproduced from Weill and Ross *IT Governance*, 2004, 11.

How Enterprises Govern

DECISION

GOVERNANCE ARCHETYPE	*IT Principles*		*IT Architecture*		*IT Infrastructure Strategies*		*Business Application Needs*		*IT Investment*	
	Input	**Decision**	**Input**	**Decision**	**Input**	**Decision**	**Input**	**Decision**	**Input**	**Decision**
Business Monarchy	0	27	0	6	0	7	1	12	1	30
IT Monarchy	1	18	20	73	10	59	0	8	0	9
Feudal	0	3	0	0	1	2	1	18	0	3
Federal	83	14	46	4	59	6	81	30	93	27
Duopoly	15	36	34	15	30	23	17	27	6	30
Anarchy	0	0	0	1	0	1	0	3	0	1
No Data or Don't Know	1	2	0	1	0	2	0	2	0	0

Most common input pattern for all enterprises.

Most common decision patterns for all enterprises.

The numbers in each cell are percentages of the 256 enterprises studied in twenty-three countries. The columns add to 100 percent.

TABLE 3: HOW ENTERPRISES GOVERN[112]

5.3 NOT-FOR-PROFIT GOVERNANCE

Weill and Ross found many similarities between non-profit and corporate IT governance, but they note that different governance approaches originate from a broader definition of public value and a culture that values consensus, transparency, and equity in not-for-profit organizations. Challenges faced by not-for-profits include difficulties in measuring performance and value, a culture of formal committees, limited budgets, endless opportunities to create value, and the fact that performance measures are not as clear as in for-profit firms.[113]

Weill and Ross recommend forming governance structures around a value framework based on authorizing environment (potential clients or customers, funding sources and political influences), capabilities (resources in the form of money and authority to create capabilities), and public value (generated in addition to private value, includes public goods and equity, and extends beyond the organization). Public value created by the not-for-profit goes back to the authorizing environment

[112] Adapted and modified from MIT Sloan School Center for Information Systems Research (CISR), 2003, Reproduced from Weill and Ross *IT Governance*, 2004, 64.
[113] Weill and Ross*, IT Governance*, 195 - 200

in the form of transfers and provision of services, and the authorizing environment grants permission back to the creators of public values in the form of market signals (increased demand) and political acceptance.

Weill and Ross argue that the authorizing framework, capabilities, and public value must be aligned toward the goal of creating public value to the maximum extent possible given the authorizing environment expectations and available internal and external capabilities.

5.4 OPEN SOURCE SOFTWARE GOVERNANCE

Benevolent Dictatorship and Meritocracy are two open source software governance archetypes.[114] In a benevolent dictatorship, a single individual or organization maintains centralized control. Project founders retain control throughout the life of the project, determine strategic direction and have the authority to make decisions, particularly when the community is in dispute. Community discussion and debate should inform the decision-making process, and decisions should be made in the best interests of the organization.

In a meritocracy, distributed control is awarded in recognition of contributions to the project, and reward mechanisms are built into the governance framework. Project leaders have earned the respect and recognition of the community, and everyone can contribute to the decision-making process. Structured decision-making processes are required, because no single individual can break a community-deadlock. Examples of open source software governance models can be found below (Table 4).

Common Open Source Software Projects and Their Governance Framework

Project type	Objective	Control style	Community structure	Examples
Exploration-oriented	Sharing innovation and knowledge	Cathedral-like central control	Project leader, many readers	Perl, Linux Kernel
Utility-oriented	Satisfying an individual need	Bazaar-like decentralized control	Many peripheral developers, peer support to passive users	GNU/Linux (excluding Kernel)
Service-oriented	Providing stable service	Council-like central control	Core members instead of project leader, many users that develop systems for end users	Apache, PostgreSQL

TABLE 4: COMMON OPEN SOFTWARE GOVERNANCE MODELS[115]

[114] Ross Gardler and Gabriel Hanganu, "Governance Models."
[115] Ibid.

5.5 GOVERNANCE MECHANISMS

Governance is implemented on a day-to-day basis via governance mechanisms, which include decision-making structures, alignment processes, and communication processes.[116] Weill and Ross suggest effective governance mechanisms must be aligned with the organization's strategy, structure, and desired outcomes.

Decision-Making Structures: Organizational units and roles responsible for making IT decisions, including committees, executive teams, and business/IT relationship managers.

Alignment Processes: Formal processes that ensure daily behaviors are consistent with IT policies, and provide input for IT decisions. They include IT investment proposal and evaluation processes, architecture exception processes, service-level agreements, chargeback, and metrics.

Communication Processes: Announcements, advocate, channels, and education efforts that inform the organization of IT governance principles and policies, and outcomes of IT decision-making processes.

Effective governance mechanisms should be simple, transparent, and suitable, and should "unambiguously define the responsibility or objective for a specific person or group."[117] Individuals should be engaged in the best way to make given decisions, and individuals affected by, or who want to challenge, the governance process should be aware of how the mechanisms work.

5.6 CHARACTERISTICS OF TOP GOVERNANCE PERFORMERS

Weill and Ross identified common qualities among governance top-performers. Top-performing firms had governance structures linked to performance measure on which they excelled (for example, growth or return on assets). Top performers also aligned business objectives with governance approaches, governance mechanisms, and performance goals and metrics. Top governance performers had seven characteristics in common: [118]

1. More managers in leadership positions can describe IT governance
2. Senior management is engaged in governance and communicates to the rest of the organization
3. More direct involvement of senior leaders in IT governance
4. Clear business objectives for IT investment
5. Differentiated business strategies
6. Fewer renegade and more formally approved exceptions
7. Fewer changes in governance from year to year

Weill and Ross also attributed effective governance performers with the specific archetypes used for gathering input and making decisions. Top governance performers used the Federal model to gather input about IT principles and business application needs, but used the Duopoly model to actually make decisions (Table 5).

[116] Weill and Ross, *IT Governance*, 86
[117] Ibid., 114-115.
[118] Weill and Ross, *IT Governance,* 145 – 146.

Best and Worst Governance Performers Use Different Arrangements

DECISION

GOVERNANCE ARCHETYPE	*IT Principles*		*IT Architecture*		*IT Infrastructure Strategies*		*Business Application Needs*		*IT Investment*	
	Input	**Decision**	**Input**	**Decision**	**Input**	**Decision**	**Input**	**Decision**	**Input**	**Decision**
Business Monarchy										
IT Monarchy										
Feudal								-		
Federal	+	-		-		-	+			
Duopoly	-	+					-			+
Anarchy										

- Poor performers + Top performers

TABLE 5: BEST AND WORST GOVERNANCE PERFORMERS USE DIFFERENT ARRANGEMENTS[119]

Successful governance in not-for-profits includes partnerships and joint decisions between IT and not-for-profit leaders, and heavy use of mechanisms, such as formal committees, and a broad definition of value (recognized by the inclusion of representatives from outside the not-for-profit in IT governance mechanisms). Stable governance mechanisms that change with less frequency are also a key characteristic, because communication and implementation of new procedures takes longer than in for-profit firms. Weill and Ross offer a number of suggestions for successful not-for-profit governance:[120]

1. Executive committees focused on all key assets including IT
2. IT council comprising business and IT executives
3. IT leadership committee comprising IT executives
4. Architecture committee
5. Tracking of IT projects and resources consumed
6. Business/IT relationship managers

[119] MIT Sloan School Center for Information Systems Research (CISR), 2003. reproduced from Weill and Ross, *IT Governance*, 131.
[120] Weill and Ross, *IT Governance, 205-206*

5.7 POTENTIAL BARRIERS TO ADOPTING A GOVERNANCE MODEL

There are a number of commonly cited barriers to adopting a governance framework.[121] These barriers include the governance process being perceived as 'red tape,' concerns that a governance framework may cause the project to lose its sense of direction, fears that the control of the project will be lost, and a belief that the project is too young or too small to attract active users or developers.[122] The first three concerns are valid, but can be remediated with a proper governance model. The fourth concern (belief that the project is too young or too small to warrant a governance model) is never valid, because potential contributors will always need clear guidance about how their contributions will be handled, and projects cannot afford to lose contributors at the beginning. These concerns should be addressed at the beginning of the project to avoid becoming a barrier to adopting a governance model later on.[123]

5.8 HOW TO CREATE AN OPEN SOURCE SOFTWARE GOVERNANCE CHARTER

Open Source Software Watch recommends that the governance charter, or terms of reference, should outline decision-making processes and explain how contributions will be handled. The initial governance framework does not have to be very detailed, and every potential situation that may arise in the future does not have to be accounted for. Instead, the model should be easily understandable and manageable, and should send a clear message to potential project contributors. The document should be concise, accessible, and easy to refer to, thus ensuring that it is used throughout the lifespan of the project.[124] Although these guidelines are focused specifically on open source software governance, many of the recommendations are pertinent governance of any type of entity.

Overview. An overview of the governance charter should include a brief summary of project objectives, how the project will be managed, and how potential contributors can become involved. It should also identify the copyright holder(s), and explain the licenses covering project outputs.

Roles and Responsibilities. The formal roles and authority within the project should be established in the governance document, specifying the level of commitment required and how individuals can take on different roles. These roles can be general or specific, and should strike a balance between too many specific roles (which would limit contributions), and not enough direction (which can lead to a lack of contribution).

Support. Support processes within the project (engagement of end-users, for example) should be defined, including available means of support, and how each one may be used. This will prevent spreading support too thin, which may limit the ability of the community to be self-supportive.

---

[121] Ross Gardler and Gabriel Hanganu, "Governance Models
[122] Ross Gardler and Gabriel Hanganu, "Governance Models
[123] Ibid. Ross Gardler and Gabriel Hanganu, "Governance Models."
[124] Ibid.

Decision-Making Processes. Openness and transparency in the decision-making processes should be emphasized. There should be a focus on striking a balance between decision-making processes being too time-consuming and not being well-defined.

5.9 PERTINENT ISSUES TO GOVERNANCE

5.9.1 LESSONS FROM HISTORICAL INFRASTRUCTURE DEVELOPMENT

There are several common patterns, best practices, and lessons learned from past cyberinfrastructure projects that are pertinent to EarthCube. If EarthCube is to be a distributed cyberinfrastructure as NSF envisions it, the EarthCube governance framework will need to consider how past infrastructures have developed.

According to "Understanding Infrastructure: Dynamics, Tensions, and Design - Report of a Workshop on 'History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures'" compiled by Edwards, Jackson, Bowker and Knobel, infrastructures, such as radio, television, electricity, the telephone and the internet, for example, followed a similar pattern as they developed.

Cyberinfrastructures are composed of local, global, social, and technical elements and are defined as "the set of organizational practices, technical and infrastructure, and social norms that collectively provide for the smooth operation of scientific work at a distance. All three are objects of design and engineering; a cyberinfrastructure will fail if anyone is ignored."[125] Edwards et al. argue that cyberinfrastructure development is a series of choices between a social and/or technical solution for any given problem, or some combination of the two (Figure 6).

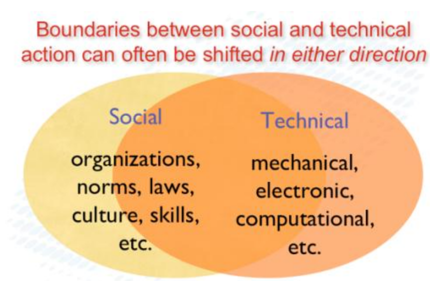

FIGURE 6: SOCIAL AND TECHNICAL BOUNDARIES[126]

Edwards et al. found that infrastructure building is usually done by many builders, planning is not always intentional, progress is usually incremental and modular, and the final version of the infrastructure is usually very different from the initial version. They also found that science, theory, and inquiry are created locally, and then brought in as necessary as the infrastructure grows and new communities contribute to the development process. Edwards et al. argue that it is

[125] Edwards et al., "Understanding Infrastructure," 6.
[126] Ibid., 4.

necessary to resolve social, cultural, and organizational issues in tandem with IT services in order to create a strong and lasting infrastructure.

Edwards et al. identified **three common steps** in historical infrastructure development:[127]

1. **System building:** Systems are linked into networks, leading to distributed control and access
2. **Technology Transfer:** Technology is transferred from one location or domain to another and the system adapts to new conditions
3. **Integration and Consolidation:** Heterogeneous systems are linked into networks through gateways, consolidating the infrastructure project

They also identified **four common infrastructure dynamics**[128]

1. **Reverse Salients:** These are critical, unsolved problems, defined as, "points where an advancing front is held back," and can be technical, social, cultural, legal, political and/or organizational.
2. **Gateways.** These are the "plugs and sockets that allow new networks to be joined with existing networks with minimal constraint." They are usually the joining together of a technical solution and a social choice, such as standards, that are integrated into users' communities of practice.
3. **Path Dependence.** Infrastructures are very difficult to alter once they are established. Path dependence occurs when early choices constrain available options later.
4. **Scale Effects:** It takes 40-50 years for short-term, local infrastructure projects to grow into large-scale, functional infrastructures.

Edwards et al. argue that infrastructure development often faces a series of **tensions** that challenge the idea that infrastructure building is planned, orderly, or mechanical.[129]

1. **Time, Scale, and Agency.** Time (short-term funding decisions versus the long-term time-scale needed for infrastructures to grow); Scale (choices between worldwide interoperability and local optimization); Agency (how to navigate planned versus emergent change).
2. **Winners and Losers.** Those who benefit from certain decisions or the course of infrastructure development, and those who are left behind.
3. **Interest and Exclusion.** The unequal distribution of power and consequences of infrastructural change among stakeholders and contributors.
4. **Ownership and Investment.** Tensions of ownership and investment occur when changing infrastructures run into intellectual property rights, public/private investment models, and policy objectives.
5. **Data Storage, Preservation and Effective Curation.** Data storage, preservation and curation are at the center of some of the most complex challenges to cyberinfrastructure development.
6. **Data Meaning.** Data have different meanings to different people.

---

[127] Ibid., 7-14.
[128] Ibid., 14-19.
[129] Edwards et al., "Understanding Infrastructure," 24-33.

7. **Data Sharing.** Data sharing necessitates tools that have the ability to accommodate and translate different needs from distinct domain communities.
8. **Trust.** Data trust is a key component of data sharing and involves data generators and data users.

Edwards et al. offer several **design recommendations** to mitigate the tensions described above:[130]

1. **Purpose.** Cyberinfrastructure design should permit collaborative work that enables scientists from different domains to work together
2. **Boundary Work.** Cyberinfrastructure can bridge the 'great divide' between system building and social analysis
3. **Principles of Navigation.** It is necessary to establish a set of design principles of navigation instead of a rigid roadmap because there is no one way to design cyberinfrastructure.
4. **Standards and Flexibility.** Flexibility should be included in cyberinfrastructure design to avoid path dependence and facilitate change

In addition, Edwards et al. offer several **recommendations for future cyberinfrastructure development:**[131]

1. **Learn from successes and failures of past cyberinfrastructure.** Comparative social scientific analysis of cyberinfrastructure, accurate and realistic reporting mechanisms, aligning extant or creating new incentive structures, and "instrumenting" cyberinfrastructure for social scientific analysis will promote learning from past cyberinfrastructure.
2. **Improve current cyberinfrastructure analysis.** Infrastructural diagnostics, training for information managers, cross-disciplinary symposia for early-career scientists, and fitting funding to cyberinfrastructure-development time scales can improve current cyberinfrastructure analysis.
3. **Enhance cyberinfrastructure resiliency and research.** Cyberinfrastructure resiliency, sustainability, and reach can be enhanced through design flexibility, mechanisms to incorporate under-represented groups and institutions, and alliances with niche organizations to eliminate redundancies in expertise.

Edwards et al. argue that the long-term success of any cyberinfrastructure is dependent on long-term, focused attention to the issues described above. For a more detailed review of these issues and other lessons from historical infrastructures, please see Appendix 1: Research Review.

5.9.2 GEOINFORMATICS GOVERNANCE RECOMMENDATIONS

Recent geoinformatics summits and workshops identify a number of governance-related geoinformatics issues. The *Workshop on Working towards a National Geoinformatics Community*

[130] Ibid., 33-37.
[131] Edwards et. al., "Understanding Infrastructure," ii.

(NGC) (2010) identified several governance issues that plague the geoinformatics community, including problems of data integration, understanding, trust, and data longevity. [132]

The Earth and Space Informatics Summit (2008) report recommended a two-pronged approach to expanding the geoinformatics community by creating new contacts and working with existing and planned structures. Additionally, more governance structures may be added as necessary, and communication gaps should be dealt with. Other recommendations include international agreements on best practices and standards, WDC certification for relevant groups, a generic, agreed-upon ICSU (International Council for Science) data policy, shared data policy resources, identification of good ideas that are ready to be shared, and open and timely access to data for scientific research.

5.10 CASE STUDIES FROM GOVERNANCE RESEARCH REVIEW

A series of case studies from the governance research review demonstrate that governance is as varied as the institutions it serves, even within the same types of communities (see Appendix 1: Research Review).

Case studies include:

- Domain science; CUAHSI, IRIS, Unidata
- IT: W3C Consortium, DataNet Federation Consortium, DataONE, Apache Software Foundation
- Federations: ESIP and OGC
- Large Facilities: SURA US IOOS Coastal and Ocean Modeling Testbed, Project GENI

Case studies key points:

1. Each organization has a governance framework developed specifically to meet that organization's needs. For example, even within the federated organizations, OGC's governance framework differs greatly from that of ESIP.
2. Organizational goals vary greatly from organization to organization, as do the management structures and decision-making processes employed to achieve those goals.
3. Funding mechanisms and paid staff vs. volunteers differ greatly from organization to organization.
4. The effectiveness of the governance framework should be evaluated against how well the framework enables the organization to move forward in achieving its goals.

A case study from each of the four types of groups is included here. Areas of focus are 1) general background, 2) funding, 3) governance, and 4) staff.

5.10.1 DOMAIN SCIENCE: CUAHSI

General Background. The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) represents over 130 U.S. universities and international hydrologic-science

[132] Walker, J. Douglas, Allison, Lee, and Gundersen, Linda, conveners. (in preparation). Workshop on Working towards a National Geoinformatics Community (NGC), USGS Denver Federal Center, Denver, Colorado, September 23-24, 2010, 1.

organizations. Its mission is to enable "the university water science community to advance the understanding of the central role of water to life, Earth, and society,"[133] and it is composed of collection of hydrologic databases. CUAHSI supports its community to "advance water science and to improve societal well-being." [134]

Funding. The CUAHSI Hydrologic Information System (CUAHSI-HIS) is an NSF-funded project whose mission is to improve access to water data through the development of infrastructure and services.[135] Specific services include tools, standards and procedures, web services, and standards and procedures. The CUAHSI-HIS architecture is composed of three main components: data publication, data cataloging, and data discovery, access, and analysis.

Governance. The CUAHSI governance framework is guided by a set of by-laws: [136]

- **Article I: Name**
- **Article II: Member Institutions:** Describes how institutions may become CUAHSI Members or Affiliates.
- **Article III: Meetings of Membership and Election of the Board of Directors:** Provides requirements for the annual meeting, notice of meetings, quorum, nominating committee, ballot, election of directors, method of voting, counting of ballots, special elections, special meetings, and virtual participation.
- **Article IV: Board of Directors:** Describes the powers, composition, election, term of office and processes for resignation, removal, transfers and vacancies of Board Members.
- **Article V: Meetings of the Board of Directors:** Provides requirements for the annual meeting, special meetings, notice of meetings, quorum, voting, action without a meeting, and virtual participation.
- **Article VI: Officers:** Describes officer qualifications, roles of the Chair Person, President, Secretary, and Treasurer, elections and term of office, and the processes for resignation, removal and vacancies of Officers.
- **Article VII: Executive Committee of the Board, Other Committees and Advisory Council:** Describes the process for the creation of new committees, the powers of the Executive Committee, the role of the Internal Audit Committee, Special Committees, Standing Committees, Senior Advisory Council, and the Corporate Advisory Board.
- **Article VIII: Election of Officers and Executive Committee at the Annual Meeting of the Board of Directors:** Describes the nominating and voting processes to elect officers and Executive Committee members at the annual meeting.
- **Article IX: Fees and Assessments:** Requires that new members contribute an initiation fee and additional fees as determined by majority vote of the Board of Directors.
- **Article X: Compensation:**

[133] "What is CUAHSI?" Consortium of Universities for the Advancement of Hydrologic Science, Inc., accessed June 1, 2012, http://www.cuahsi.org/docs/What-is-CUAHSI.pdf.
[134] Ibid.. What is CUAHSI?" Consortium of Universities for the Advancement of Hydrologic Science, Inc., accessed June 1, 2012.
[135] Ibid.
[136] See "Bylaws of the Consortium of Universities for the Advancement of Hydrologic Science, Inc.," cuahsi.org, adopted December 7, 2010, http://www.cuahsi.org/by-laws.html.

- **Article XI: Indemnification**
- **Article XII: Fiscal Year**
- **Article XIII: Seal of the Corporation**
- **Article XIV: Amendments to the Bylaws:** The CUAHSI bylaws may be amended by an affirmative vote of 60% of the Members, Non-profit Affiliates and International Affiliates.

Staff. CUAHSI is governed by the Board of Directors and managed by the Executive Committee (composed of Board members and Officers). CUAHSI Committees and a small paid staff carry out CUAHSI activities (Figure 7).

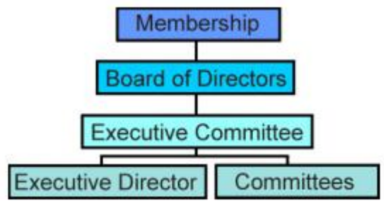

FIGURE 7: CUAHSI LEADERSHIP STRUCTURE[137]

5.10.2 IT: DATANET FEDERATION CONSORTIUM

General Background.[138] The DataNet Federation Consortium is an NSF funded project to build national data cyberinfrastructure for NSF funded research projects. Current participants include Ocean Observatories Initiative, Hydrology projects, and the iPlant Collaborative. DataNet expects policies to be developed independently by each project for governance, management, administration, and assessment. The integrated Rule Oriented Data System (iRODS) allows each group to enforce their own policies, independently of the other projects. The infrastructure is designed to be highly extensible, separating policy enforcement from data manipulation, data storage, and data access mechanisms.

While the software infrastructure is generic, the policies and procedures can be unique to each user community. Applications of the software include a DataNet Federation Consortium grant, projects with the iPlant Collaborative, the Ocean Observatories Initiative, a LifeTime Library student digital library at SILS, a collaboration with RENCI on E-iRODS development, and more.

The software is open source. There is an active international community who provide extensions to the software. The contributors include DistributedBIO, IN2P3 (France), University of Liverpool, Academia Sinica (Taiwan), ARCS (Australia), LSU Peashare, and more. Collectively across the international collaborators, 50 persons typically attend a yearly iRODS user group meeting; the number of users of the technology is much larger.

Funding. DataNet Federation is funded by NSF.

---

[137] "Organization," Consortium of Universities for the Advancement of Hydrologic Science, Inc., accessed June 1, 2012, http://www.cuahsi.org/organization.html.
[138] DataNet Federation case study written by Reagan Moore, personal emails, May 31, 2012 and June 7, 2012.

Governance. The iRODS data grid governance framework is mapped below (Table 6).

Decision	*IT Principles*	*IT Architecture*	*IT Infrastructure Strategies*	*Business Application Needs*	*IT Investment*
Archetype					
Business Monarchy					
IT Monarchy					
Feudal	X	X		X	
Federal			X		X
Duopoly					
Anarchy					

TABLE 6: IRODS DATA GRID GOVERNANCE FRAMEWORK.[139]

1. **What decisions must be made to ensure effective management and use of IT?**
   We think in terms of the driving purpose for the shared collection.
   a. The purpose is translated into a set of properties that should be enforced over time.
   b. The properties are enforced through management policies, which are cast as computer executable rules.
   c. The rules control the execution of procedures, which generate the desired properties.
   d. The procedures create state information which is automatically saved.
   e. The state information can be queried by periodic policies to verify that the desired properties have been maintained.
   f. This automates management and administrative tasks, and provides a way to validate the system. We strongly believe that automation of policy enforcement is an essential component of large-scale data management systems.
2. **Who should make these decisions?**
   a. The community that is promoting the shared environment should make the decisions. The policies can be augmented by requirements from funding organizations (data management plans), legal requirements (HIPAA, FERPA), and institutional commitments (resources).
3. **How will these decisions be made and monitored?**
   a. The consensus is the difficult step. Fortunately, policy-based data management systems support the evolution of the policy data sets. If the initial policy is insufficient, it can be updated without having to modify code. Versions of policies can be applied, and collections can be migrated from an original set of policies to a new set of policies.

Staff

The number of persons in the DICE Center who work on iRODS is 9. The number of persons at RENCI who work on iRODS is 5. The total number of paid staff is 14. Note that some of these persons are part time.

[139] Reagan Moore, personal email, May 31, 2012.

5.10.3 FEDERATED GROUPS: ESIP, OGC

ESIP and OGC are polar opposites within federated groups, so both case studies are included as examples of the wide variety of existing governance frameworks.

FEDERATION OF EARTH SCIENCE INFORMATION PARTNERS (ESIP)

Background information. [140] The Federation of Earth Science Information Partners (ESIP) is a broad-based, distributed community of science, data and information technology practitioners that coordinate interoperability efforts across the Earth and environmental science communities. Participation in the ESIP Federation is beneficial to individual members by providing an intellectual commons to expose, gather and enhance their own in-house capabilities in support of their organization's mandate. By virtue of working in the community, ESIP members experience the network effect, which enables more coordinated interoperability efforts across domain-specific communities. The ESIP Federation has a 14-year track record of success and continued growth using this community-based, discipline and agency neutral approach. These efforts catalyze connections across organizations, people, systems and data allowing for improved interoperability in distributed systems. The ESIP Federation is working in areas at the leading edge of technology development including: data architecture, data management and preservation, tool creation, cloud computing, semantic web, data systems integration and data management professional development.

Initially created by NASA in 1998, the ESIP Federation was formed in response to a National Research Council recommendation calling for the involvement of community stakeholders in the development of NASA's EOSDIS as a critical element of the U.S. Global Change Research Program (http://www.gcrio.org/USGCRP/LaJolla/cover.html). Since its inception, the ESIP Federation has continually grown and attracted a diverse group of partners, now including more than 135 member organizations. The ESIP Federation's membership is broad, including federal data centers, government research laboratories, research universities, education resource providers, technology developers, and various nonprofit and commercial enterprises. This diversity (multi-sector, agency, and function) provides expert capabilities on which participating organizations can draw.

Funding. The ESIP Federation is principally supported by NASA and NOAA.

Governance Structure. The ESIP Federation is governed by its Assembly, the standing body which affords each partner one vote. Governance documents include a Constitution and Bylaws as well as a 5-year strategic plan (http://wiki.esipfed.org/index.php/Federation_Documents). In between annual Assembly business meetings, the ESIP Federation is governed by its Executive Committee, a body comprised of the ESIP Federation's elected President and Vice President, its Standing and Administrative Committee chairs and by members representing each ESIP type (Data Center, Research, Application Developer). All Executive Committee members are either elected by the Assembly or by their ESIP Type caucus. While the ESIP Federation remains unincorporated, the Foundation for Earth Science (a 501(c) 3) provides financial and other management services to the ESIP Federation.

---

[140] ESIP case study written by Carol Meyer of ESIP, personal email, June 5, 2012.

Staff. The ESIP Federation runs with support from two professional staff and a host of volunteers. contributions from experts across the Earth science data and technology community.

OPEN GEOSPATIAL CONSORTIUM (OGC)

General Background.[141] OGC is a Voluntary Consensus Standards Organization founded 1994. It has 445 members, 38 adopted standards, hundreds of product implementations, broad user community implementation worldwide, and has alliance partnerships with 30+ standards and professional organizations. OGC is member-driven. OGC's Approach for Advancing Interoperability includes the Interoperability Program (IP), Standards Program, Compliance Program, and Marketing and Communications Program.

Funding. OGC funding comes from membership dues and project fees (Interoperability Program initiatives).[142]

Governance. OGC's Interoperability Program, Standards Program, Compliance & Interoperability Testing & Evaluation (CITE) Program, and Marketing and Communications are governed by a Board of Directors, Strategic Member Advisory Committee, Global Advisory Council, and Executive Director. Each program then has its own governance framework and mechanisms designed to carry out that program's goals. For example, "In the OGC Standards Program the Technical Committee and Planning Committee work in a formal consensus process to arrive at approved (or "adopted") OGC® standards,"[143] while the "Interoperability Program is a series of hands-on engineering initiatives to accelerate the development and acceptance of OGC standards." [144] OGC's governance framework is represented in Figure 8 below.

---

[141] David Arctur of OGC contributed to the writing of the OGC case study.
[142] David Arctur, personal email, June 6, 2012.
[143] "Programs," Open Geospatial Consortium, last modified 2012, http://www.opengeospatial.org/ogc/programs.html.
[144] Ibid...

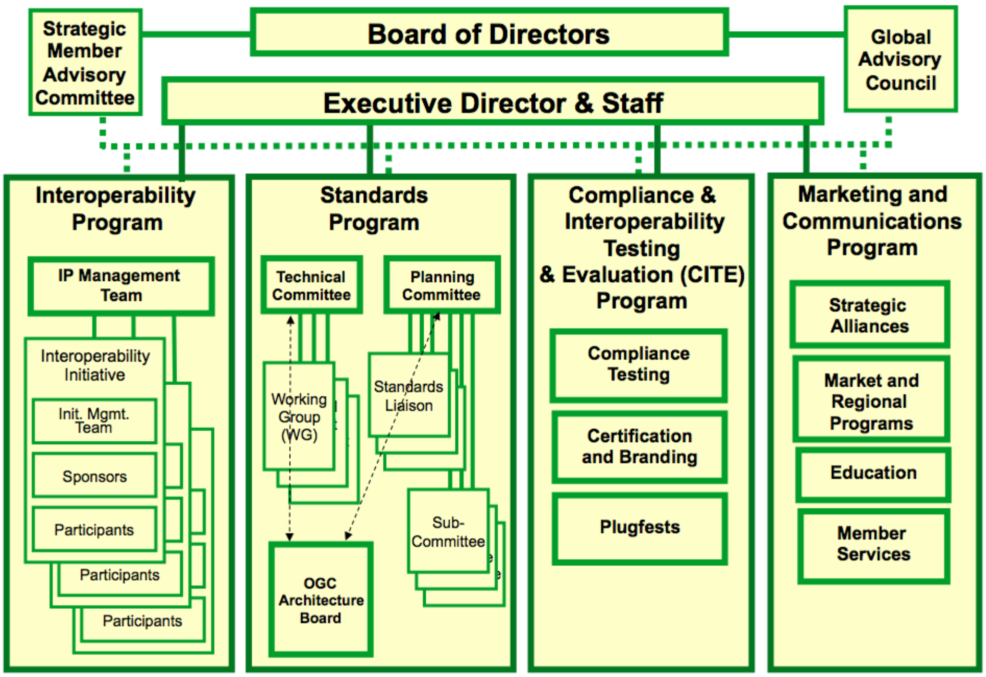

FIGURE 8: OGC ORGANIZATION AND PROGRAMS

The basic model for standards development is for a group of OGC members to submit a proposal to start a Standards Working Group for the specific technology and version (e.g., OGC Sensor Observation Service version 2.0). The group forms, subject to review by the OGC Architecture Board and a vote by the OGC members. The SWG (Standards Working Group) does its work, and when they are ready (a few months to a few years later) the completed specification documents are submitted for adoption vote by the OGC members.

An alternate path is also possible when the standards development effort is done outside OGC. In this case, the specification documents are normally be reformatted to be more consistent with OGC–native standards, and the proposal is be presented as a Request For Comment (RFC) to the OGC members, following the OGC Architecture Board's review. The most notable example of this being done was to get OGC branding for CF/NetCDF, which is developed outside the OGC.

OGC has a set of Policies and Procedures to fulfill its mission and vision:[145]

1. Current Bylaws
2. Principles of Conduct
3. Technical Committee Policies and Procedures document for Standards Development.
4. Legal Notices & Term and Conditions
5. Intellectual Property Rights (IPR)
6. Interoperability Program Policies, Processes, and Procedures (IP PP&P)
7. Policies & Procedures for Translation of OGC Public Documents
8. Policies & Procedures for the OGC Architecture Board (OAB)
9. Policies & Procedures for the Compliance Testing Program (CITE)

[145] "Policies," Open Geospatial Consortium, last modified 2012, http://www.opengeospatial.org/ogc/policies/.

Staff. "Director, Interoperability Programs" are the program & project managers for Interoperability Program initiatives. OGC staff cannot be working group or subcommittee chairs. At times OGC hires a consultant to lead an interoperability initiative, depending on resourcing and expertise available. These are drawn from a pre-qualified group called the IP Pool. (http://www.opengeospatial.org/projects/ippool). Some are contributed by member companies; some are paid by project funds.[146]

5.10.4 LARGE FACILITY: SURA/US IOOS COASTAL AND OCEAN MODELING TESTBED

General Background.[147] The key elements of the Southeastern University Research Association/ U.S. Integrated Ocean Observing System (SURA U.S. IOOS) Super-Regional Modeling Testbed governance that have been crucial to success are:

1. The appointment of a non-conflicted Testbed Advisory and Evaluation Group (TAEG) to provide objective, independent guidance, and assessment of progress;
2. Distributed leadership: each team had a team lead and held regular telecons/meetings of all participants;
3. Web site to enable knowledge management and facilitate communication;
4. Centralized management and project coordination by SURA.

Funding. The SURA U.S. IOOS Super-Regional Modeling Testbed is funded by the National Oceanic and Atmospheric Administration (NOAA).

Governance. SURA's overall management and project coordination for the Testbed not only ensured that the federal funds were expended in a responsible and accountable manner, but also includes the full support by SURA's financial and accounting services, grants and contracts management services, and support services. The Testbed enjoyed the full endorsement of the SURA Coastal and Environmental Research Committee (CERC) which is chaired by Dr. Chris D'Elia of LSU. Although not exercising day-to-day project management, the CERC is the "authorizing" entity for SURA Coastal Research projects.

The Testbed Advisory and Evaluation Group (TAEG) is an independent, technical and scientific advisory group comprised of scientific and computer experts from academia and the federal government. The TAEG makes objective recommendations regarding team selection, resource allocations and progress assessments and provide a long- range vision and Science Plan for the project. The Testbed Principal Investigator (PI) works closely with the SURA Project Director, who in turn reports to the SURA CEO. The Project Coordinator at SURA is responsible for the day-to-day activities and communications among Testbed PIs and teams, and also is the liaison among the SURA HQ, the U.S. IOOS Program Office and the Testbed PIs and teams (Figure 9).

[146] David Arctur, personal email, June 6, 2012.
[147] SURA case study written by Gary Crane of SURA, personal email, June 5, 2012.

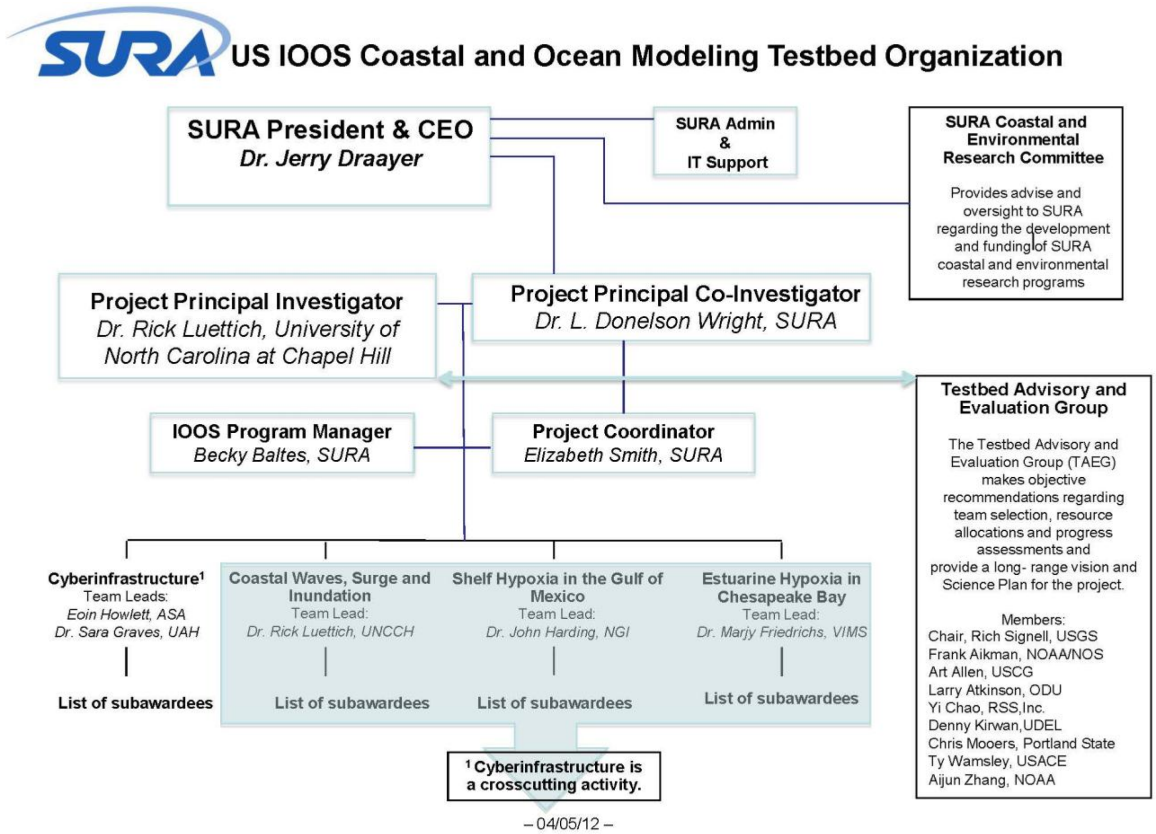

FIGURE 9: SURA US IOOS COASTAL AND OCEAN MODELING TESTBED ORGANIZATION[148]

The Testbed is comprised of three modeling teams (Inundation, Estuarine Hypoxia and Shelf Hypoxia) each with a team lead, and a cyberinfrastructure team which is developing the overarching, and sustaining Testbed framework which will be for the benefit of the modeling teams and others wishing to utilize the framework.

Staff. The Coastal and Environmental Research Committee and the Testbed Advisory and Evaluation Group are both volunteer groups. All other participants (with the exception of the IOOS Program Manager) are supported directly or indirectly through external funding.

[148] Courtesy of Gary Crane, personal email, Jun 5, 2012.

5.10.5 FUNDING MODELS

The following table summarizes the funding models of various domain science, IT, federated, and large facilities groups, some of which were reviewed previously as case studies (Table 7). Funding mechanisms of the governance framework will help determine the final structure; however, that structure must remain agile enough to adapt when, and if, a funding stream is exhausted.

Category	Program	Governance Funding Model
Domain Science	Unidata, IRIS, IEDA	Core Grant
	UCAR, CUAHSI	Membership Fee
IT Groups	W3C	Membership Fee
	Apache	Foundation(donations)
Federated Groups	OGC ESIP Federation	Membership Fee, Grants for Interoperability projects *Grants*
	DataONE	Core Grant
Large Facilities	OOI, EarthScope, NEON, XSEDE, IOOS	Core Grant

TABLE 7: SUSTAINABILITY/FUNDING MODELS FOR DOMAIN SCIENCE, IT GROUPS, HYBRID GROUPS, AND LARGE FACILITIES.[149]

5.10.6 THE ELECTRICAL GRID AS AN EXAMPLE OF A SYSTEM OF SYSTEMS MODEL

Most of the case studies examined for EarthCube come from organizations engaged in some aspects of geoscience cyberinfrastructure, and this is true for the case studies reviewed in the Governance Roadmap as well. We looked at OGC and ESIP for example, as models for standards adoption and community building respectively. We look at IRIS's Data Management Center and UCAR's Unidata as models (see Appendix 1: Research Review) for data management. These case studies focus on single entities however, and not a complex system of linked entities that is typical of most infrastructure and cyberinfrastructure.

The electrical grid illustrates the level of complexity, independent, dependent, and interdependent relationships, and range of functional options for many pieces that are comparable to the nature of building and deploying cyberinfrastructure. There is no single entity responsible for the U.S. electrical grid. It is considered one of the greatest engineering achievements of the 20th century and is remarkable for its pervasive deployment and phenomenal reliability. It is a type case of a successful infrastructure (Figure 10).

[149] Mohan Ramamurthy, "Governance Examples" (presentation, EarthCube Governance Virtual Plenary Workshop), April 11 and 17, 2012.

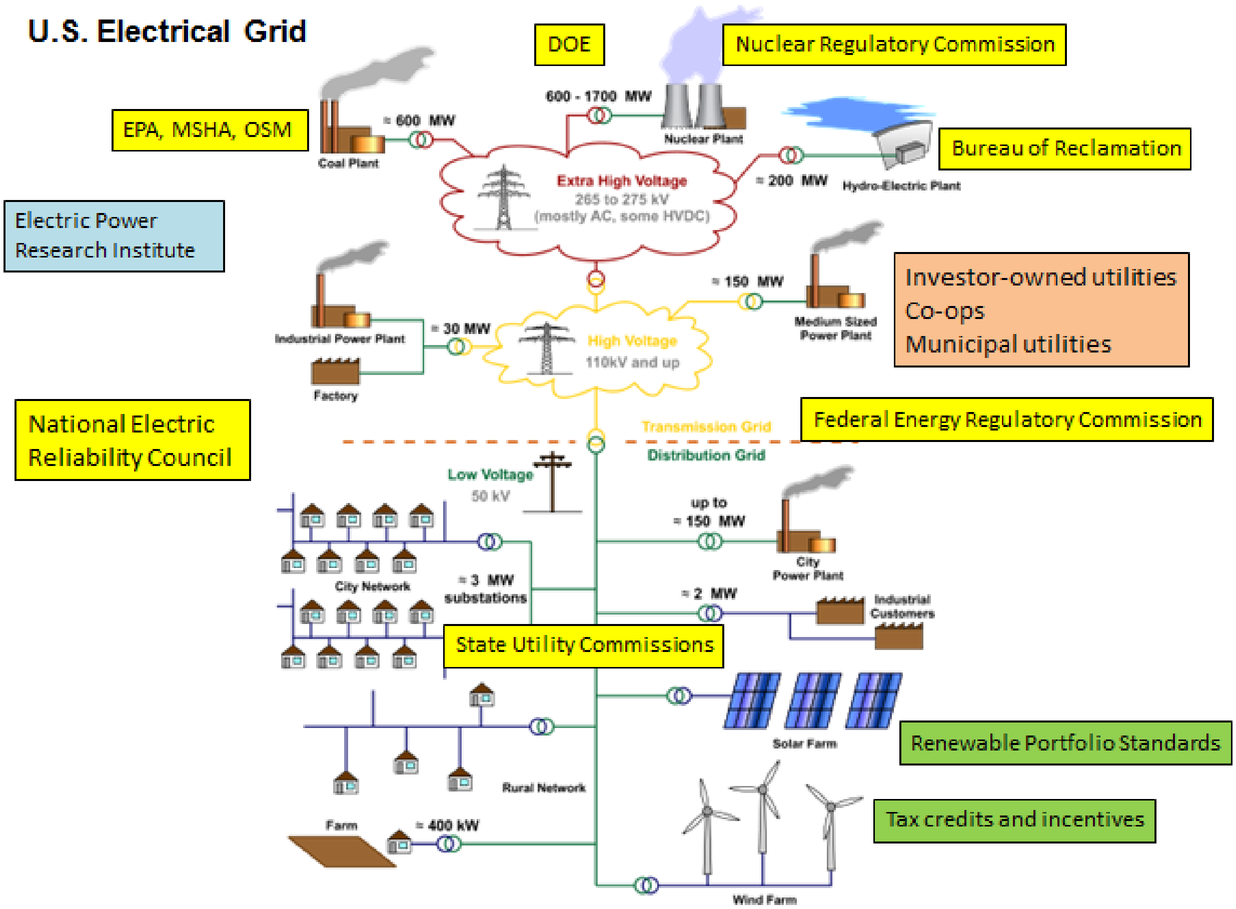

FIGURE 10: THE U.S. ELECTRIC GRID OVERLAPPING AND INTERLEAVED MANAGEMENT AND GOVERNANCE SYSTEMS, WITH ANNOTATIONS FROM THIS STUDY[150]

Take for instance, electric utilities. All serve the purpose of delivering electricity to their customers, reliably, economically, and in a consistent format. However, there are multiple types of organizations that carry out these functions under different management and organizational models, and under different governing or regulatory systems. Investor-owned utilities are profit-making and typically overseen by state utility commissions that set rates and approve the types of power generation technology used. Electric co-operatives however, are owned by their customers and generally exempt from the same utility commission oversight, as are municipal utilities that report to voters.

Yet, every utility, regardless of their organization and internal governance, is subject to governance by federal and state agencies, based on the type of power plants they operate. Thus, nuclear power plants are overseen by the Nuclear Regulatory Commission (NRC) whether they are operated by an investor-owned utility or a co-op.

[150] Base diagram from Wikipedia, "Electrical Grid," http://en.wikipedia.org/wiki/Electrical_grid, accessed May 30, 2012.

The Federal Energy Regulatory Commission (FERC) is engaged in approval of the size, location, and interconnections of electric transmission. The FERC-authorized North American Electrical Reliability Corporation (NERC) serves as the Electrical Reliability Organization for the U.S. NERC's policies have evolved into standards that can be imposed on interconnected regional power systems to ensure reliability. But FERC operates independently of NRC, utility commissions, and utilities.

Running concurrently, and generally non-overlapping, to these operations are agencies that deal with production, safety, pollution, and disposal of power generating fuels (i.e., coal, nuclear, gas, hydro). Superimposed over this are tax policies and financial incentives for alternative power technologies, such as wind, solar, and geothermal.

A key concept to take away from the U.S. Electrical Grid is that something this complex and important, both technically and organizationally, cannot be governed by one single model or one encompassing entity. In our research, we have not found an infrastructure that is governed by a single entity let alone a single governance archetype.

5.11 GOVERNANCE RECOMMENDATIONS FROM EARTHCUBE WHITE PAPERS

EarthCube White Papers (submitted fall 2011, Governance Category and Designs Category) contain several recommendations for EarthCube governance (see Appendix 1: Research Review for more information).

5.11.1 THEMATIC RECOMMENDATIONS

Clarity. EarthCube governance should make clear how components should be managed, how EarthCube operates with infrastructures, and how EarthCube will evolve.

Collaboration. Governance must promote the collaboration of users, processes, and resources inside and outside of the geosciences. [151]

Commodity Governance and Communities of Communities. EarthCube will be composed community of communities (entities that have national and international cooperation, distributed resources, multiple tools, standards, agencies, organizational missions, goals and target communities, and no single project manager). Communities of communities require a re-envisioning of governance, promotion, collaboration and communication. [152]

---

[151] Aaron Braeckel, Bob Barron, Arnaud Dumont, and Bruce Carmichael, "EarthCube Design Approach," National Science Foundation EarthCube White Paper: Design Category, 2011.
[152] Cecelia DeLuca, "Commodity Governance for Communities of Communities," National Science Foundation EarthCube White Paper: Governance Category, 2011.

Community. The EarthCube audience must be specified and have a high degree of community buy-in.[153] EarthCube governance must advance the interests of its intended users.[154]

Community-Based Discovery Model. In this model, "eventually community-supported standards will emerge via collaborations."[155] EarthCube development processes should be transparent, responsive to community input, serve as teaching tools, and should be based on community feedback. [156]

Communities of Interest. EarthCube will be formed by Communities of Interest (COI - groups of users who share a common interest within EarthCube).[157]

Communities of Practice: Informaticists, Data Generators, Synthesizers. The geosciences community can be divided into three main communities of practice: Informaticists (scientists and computer scientists with IT and knowledge management expertise), Data Generators (researchers that collect and use observational and experimental data), and Synthesizers (community of scientists exploring research questions requiring data synthesis and analysis). They synthesizers should be engaged first.[158]

Community-Based Governance.[159] A community-based governance model is the best model for EarthCube because individual members who have a greater stake in the community as a whole are more likely to contribute to the advancement of the community, and be involved in the decision-making process. Key questions include:

- How is community created and when should this process begin?
- What does it mean to be a peer (i.e. if someone is considered a peer, they should have equal access to information and decision-making)?
- When and how is feedback collected?
- What processes are in place to respond to criticism and feedback?

Diverse Representation in Governance. There should be diverse, multi-sector participation with comprehensive expertise in the represented areas. Governance should be driven by, and be responsive to, the community needs and wants.[160]

---

[153] DavidDave Fulker, "EarthCube Leadership." National Science Foundation EarthCube White Paper: Governance Category, 2011.
[154] Tim Ahern and Keith Koper, "Flexibility in an EarthCube Design Approach: Ideas from IRIS and its Partners," National Science Foundation EarthCube White Paper: Designs Category, 2011.
[155] Jim Bowring, Samuel Bowring, and Douglas Walker, "An Earth Cube Design Approach." National Science Foundation EarthCube White Paper: Designs Category, 2011, 2.
[156] Ibid.
[157] Aaron Braeckel et al., "EarthCube Design Approach."
[158] Shanan Peters, et al., 'Towards a New Distributed Platform for Integrative Geoscience: An EarthCube Design Approach and Prototype Plan," National Science Foundation EarthCube White Paper: Design Category, 2011.
[159] Bruce Caron, "EarthCube Governance Whitepaper: Realizing expectable returns on EarthCube investments in community building and democratic governance," National Science Foundation EarthCube White Paper: Governance Category, 2011.
[160] Idaszak, Ray (ed.), et al., "Insights and Precepts from a Workshop on Creating a Software Institute for Environmental Observatories," National Science Foundation EarthCube White Paper: Governance Category, 2011.

Evolution. EarthCube must evolve to meet changing needs and opportunities. The user community itself is the most valuable source of innovation, yet "EarthCube must remain stable while supporting change."[161]

Flexibility. EarthCube should produce a flexible and reusable system that can be extended by future organizations which were not necessarily involved in the original design process.[162]

Form Should Follow Function. Forms of governance and management should be tightly linked to functions that implement EarthCube goals.[163]

Goals. EarthCube governance must reflect the long-term vision and scope of EarthCube, and the governance framework must suit long-term goals.[164]

Guidelines and Principles. The EarthCube governance framework should develop basic principles and guidelines for publishing geoscience data and data products to EarthCube.[165] There is also a need for EarthCube governance framework to develop and follow an agreed-upon process to coordinate viewpoints, decisions, and interoperability.[166]

Leadership. There must be a clear assignment of leadership.[167]

Open Source Software Governance and Open Science. Open source software and open science are models for EarthCube governance. Open source governance contributes knowledge to the commons while preserving constructive competition, and will enable scientists to more effectively address the Grand Challenges in research identified by NSF.[168]

Public-Private Partnership. A public-private partnership (government, academia, and industry) will effectively serve national, regional, and local decision-makers, the private sector, and the general public.[169]

Success. There must be a clear definition of success, with qualitative and quantitative measures to evaluate progress towards overall goals.[170]

---

161 Aaron Braeckel et al., "EarthCube Design Approach."
162 Shailendra Kumar, Stefan Falke, Martin Federick, and Rick Ohlemacher, "A Public Private Partnership Opportunity for NSF Earth Cube to Serve Stakeholders and Decision Makers," National Science Foundation EarthCube White Paper: Governance Category, 2011.
163 Ahern and Koper, "Flexibility in an EarthCube Design Approach."
164 Ibid.
165 Peters et al., 'Towards a New Distributed Platform for Integrative Geoscience."
166 ESIP-DC, "A Lightweight Approach to Earth Science Data Discovery."
167 Dave Fulker, "EarthCube Leadership."," National Science Foundation EarthCube White Paper: Governance Category, 2011
168 Dan Bedard, Ray Idaszak, Howard Lander, and Stan Ahalt, "What Open Source Software Development Teaches Us about Scientific Governance," National Science Foundation EarthCube White Paper: Governance Category, 2011.
169 Kumar, et al., "A Public Private Partnership Opportunity for NSF Earth Cube to Serve Stakeholders and Decision Makers."
170 Fulker, "EarthCube Leadership."

Sustainability. Sustainability of EarthCube is linked to the intrinsic value of investments in data collection and preservation. Thus, EarthCube must, "Demonstrate how sustained investments in data resources contribute to advances in knowledge and the public good."[171]

- EarthCube sustainability is also dependent on a governance framework that anticipates and allows for the evolving needs of the EarthCube community.[172]
- There must be trust in the system as a whole for EarthCube to be sustainable.[173]
- The EarthCube community must have sufficient incentives to bring together resources to improve and maintain the EarthCube system.[174]
- Sustainability can be promoted through use-cases that allow scientists to imagine how their research can be enhanced through cyberinfrastructure.
- Education and outreach to scientists and the broader EarthCube community can also promote sustainability.[175]
- Governance mechanisms should foster EarthCube sustainability. A potential mechanism is a 'Center for EarthCube Support' to organize operation strategies, short courses and workshops, social networking sites linking people and data, an Informatics Help Desk, and a Pilot Steering Committee.[176]

Trust. EarthCube design and governance should focus on foster trust of data, and governance framework must create a broad and shared chain of trust from data providers, to scientists, to end users of scientific results.[177]

5.11.2 TOPICAL AND CASE STUDY RECOMMENDATIONS

AGORA-net. AGORA-net graphically organizes knowledge based on argument maps that link arguments to one another, and provide forums for debate, comment, and counterargument. AGORA-net is a potential model for how knowledge may be mapped in EarthCube. The governance framework should be developed by the EarthCube community itself, and that governance must be open, transparent and inclusive.[178]

Centralized Database. EarthCube should be an "Open access archive of all publicly available geosciences data in the United States and eventually the world," whose long-term vision would be led by scientists.[179]

Critical Zone Observatory (CZO) Data. CZOData's focus on efficient data and knowledge integration to understand and model earth processes is relevant to EarthCube. [180]

---

171 Ahern and Koper, "Flexibility in an EarthCube Design Approach" 10.
172 Braeckel et al., "EarthCube Design Approach."
173 Ibid.
174 Mader et al., "An Architectural Approach for Sharing, Discovery and Knowledge Dissemination on EarthCube," National Science Foundation EarthCube White Paper: Designs Category, 2011.
175 Leetaru et al., "EarthCube Design Approach."
176 Peters et al., 'Towards a New Distributed Platform for Integrative Geoscience."
177 Braeckel et al., "EarthCube Design Approach."
178 Michael Hoffman, Judith Curry, and Calton Pu, "EarthAgora: A Collaborative Knowledge Management Tool." National Science Foundation EarthCube White Paper: Designs Category, 2011.
179 Leetaru et al., "EarthCube Design Approach."
180 Zaslavksy, et al., "Data Infrastructure for the Critical Zone Observatories (CZOData): an EarthCube Design Prototype." National Science Foundation EarthCube White Paper: Designs Category, 2011.

The Consortium of Universities for the Advancement of Hydrologic Science, Inc. Hydrologic Information System (CUAHSI-HIS). CUAHSI-HIS goals of data access, hydrologic observatories, hydrologic science, and hydrologic education, align well with EarthCube. Governance should focus on community process and open development. Specific mechanisms include establishing and documenting a governance model for development activities, creating tools to support community functions and interactions, active engagement of partners and contributors.[181]

DataONE. DataONE is an existing platform that allows scientists to focus on large-scale, long-term problems. DataONE employs stakeholder assessment, user scenarios, usability testing and engagement with domain scientists in use cases to evaluate cyberinfrastructure projects. [182]

EC Network. EarthCube should be an ECNetwork of interconnected nodes, much like the internet, to build connections among data, resources and disciplines. The governance framework should support distributed administration and available resources, create flexibility, promote scientific creativity, and take advantage of local expertise.[183]

EarthCube as a Double Loop Organization. In Double Loop organizations, the processes and mechanisms developed to fix certain problems simultaneously solve larger issues that created these problems in the first place. The governance framework allows stakeholders/community members to govern their own governance framework by agreeing to rules that solve common concerns related to interaction.[184]

Federation of Earth Science Information Partners (ESIP). ESIP's community-based governance model fosters collaboration and coordination of interoperability efforts across the geosciences community and is a potential model for the EarthCube governance framework.[185]

ESIP Discovery Cluster (ESIP-DC). The ESIP-DC is a volunteer-based organization that is consensus driven. The ESIP Discovery Cluster depends on open source software, and collaborates to solve problems of interoperability, distributed services, and adoption of different open standards. ESIP-DC's efforts to address these issues may be useful for EarthCube.[186]

Incorporated Research Institutions for Seismology (IRIS). IRIS's community governance model provides a framework for decision-making that preserves community oversight of IRIS facilities and is a potential model for EarthCube governance framework. IRIS's unifying structure encourages

---

[181] Tarboton, et al., "Advancing Solutions for an EarthCube Design. What can be learned from the CUAHSI HIS experience? National Science Foundation EarthCube White Paper: Designs Category, 2011.
[182] Michener, et al., "DataONE-Enabling Cyberinfrastructure for the Biological, Environmental and Earth Sciences." National Science Foundation EarthCube White Paper: Designs Category, 2011.
[183] Mader et al., "An Architectural Approach for Sharing, Discovery and Knowledge Dissemination on EarthCube," National Science Foundation EarthCube White Paper: Designs Category, 2011.
[184] Bruce Caron, "EarthCube as a Double Loop Organization," Virtual Democracy, last modified April 4, 2012, http://cybersocialstructure.org/2012/04/04/earth-cube-as-a-double-loop-organization/
[185] Chris Lenhardt, "Governance or Process: Building Community Through Collaboration for Shared Solutions" National Science Foundation EarthCube White Paper: Governance Category, 2011.
[186] ESIP-DC, "A Lightweight Approach to Earth Science Data Discovery," 2011, National Science Foundation EarthCube White Paper: Design Category, 2011.

operational efficiencies and coordination across the geosciences, and is a good model for EarthCube governance. [187]

Open Geospatial Consortium (OGC). OGC governance mechanisms promote good governance, including, "a proven framework of tools, policies, and procedures for governance of long-term, interdisciplinary and multi-sector collaboration for high-impact scientific research for societal benefit."[188]

Professional Networking Service. EarthCube should function as a professional networking service to link scientists to data, with a focus on professional relationships, protocols, formats, services and analytical tools. Governance should be community-based. [189]

QuakeSim. The earthquake science field, and QuakeSim in particular, can be a potential partner for EarthCube to help provide the cyberinfrastructure envisioned by EarthCube. QuakeSim promotes open processes, with a focus on the creation of data products.[190]

Scientific Forecasts. EarthCube should focus on scientific forecasts. Governance would be divided into three groups: Economic, Scientific, and Technical.[191]

Space-Time Location. A space-time location can serve as common ground for communication among Earth scientists. EarthCube governance framework must focus on challenges in data integration, foster sharing and interdisciplinary collaboration, and maintain geophysical science information for the future. [192]

UNAVCO. UNAVCO has created an effective community governance framework through a formal process of strategic planning and community involvement, which supports broad participation and effective oversight of UNAVCO programs.[193]

Unidata. Unidata's governance framework focuses on assessing and responding to community needs, including advocating and negotiating agreements on behalf of the community on data issues, fostering community interaction to promote sharing of data, tools, and ideas, and consensus building.[194]

---

[187] Tim Ahern, David Simpson and Brian Stump, "Guiding Principles and Best Practices from the IRIS Data Management System," National Science Foundation EarthCube White Paper: Design Category, 2011.

[188] David Arctur et al., "Governance: An OGC white paper for NSF EarthCube," National Science Foundation EarthCube White Paper: Governance Category, 2011, 1.

[189] James Long and Karen Remick, "A Data-Based Professional Networking System." National Science Foundation EarthCube White Paper: Designs Category, 2011.

[190] Fox et al., "EarthCube and EarthQuake Science," National Science Foundation EarthCube White Paper: Design Category, 2011.

[191] Philip Maechling, "Focus on Forecasts." National Science Foundation EarthCube White Paper: Designs Category, 2011.

[192] Anna Kelbert and Gary Pavlis, "Earth Sciences: The Tower of Babel." National Science Foundation EarthCube White Paper: Designs Category, 2011.

[193] William Holt and Mary Miller, "Community Governance: UNAVCO Facility Models for EarthCube Implementation," National Science Foundation EarthCube White Paper: Governance Category, 2011.

[194] Ethan Davis, Douglas Dirks, Linda Miller, and Tom Yoksas, "An EarthCube Design Process: Unidata's Perspective," National Science Foundation EarthCube White Paper: Designs Category, 2011; Ramamurthy, "Unidata Governance."

United States Geoscience Information Network (USGIN). EarthCube should leverage USGIN, whose goal is to build catalog and web services to link and allow access to holdings of geological surveys in the US, in addition to tools and applications to understand data.[195]

United States Geological Survey (USGS). Collaboration among government entities, and legislative mandates, form bases of interaction between government data agencies and EarthCube participants. The USGS is already involved in building and sustaining cyberarchitecture to support its Earth science mission.[196]

Western North America Interferometric Synthetic Aperture Radar Consortium (WInSAR) Consortium. The WInSAR governance model is an example of a low-cost and effective way to achieve both EarthCube and WInSAR goals.[197]

The synthesis of governance research, case studies, historical infrastructure development, EarthCube white paper recommendations, and community feedback led us to the conclusion that EarthCube governance will likely be much more complex than any single case study or governance framework of a single organization. Our conclusions regarding EarthCube governance as a system of systems is described in following section (6.0 Solutions).

---

[195] Allison, et al., "Geological Survey Contributions to Building Cyberinfrastructure in the Geosciences," National Science Foundation EarthCube White Paper: Design Category, 2011.
[196] Bristol et al., "Role of Government Data Agencies, Virtual Observatories, and Communities of Practice in Sustaining EarthCube," National Science Foundation EarthCube White Paper: Design Category, 2011.
[197] Fielding et al., "Challenges Facing the WInSAR Consortium from the Coming Tsunami of SAR Observations." National Science Foundation EarthCube White Paper: Governance Category, 2011.

6.0 SOLUTIONS

6.1 IMPORTANT GOVERNANCE ISSUES

The goal of the roadmapping process was not to select a governance model in and of itself, but to provide sufficient background research on governance, and feedback from community engagement to forge a path to move EarthCube governance forward. However, this roadmap does layout several action items and it is our hope that the implementation of the Governance Roadmap will result in the production of an initial governance framework for EarthCube (see the Milestones and Tasks section).

We recognize that EarthCube's governance is likely to evolve as EarthCube grows and develops, particularly when transitioning from the build phase to the implementation phases in EarthCube development.

Initially, EarthCube's governance must be able to meet the requirements described in Section 4: Requirements, and it will likely employ a series of governance models to address different services and functions across EarthCube. It is possible that EarthCube will initially require a more centralized governance model, but as it grows into a system of systems, it is likely that a more distributed governance framework will be most appropriate.

Thus, in addition to crafting the initial governance framework, it is important to take into account how that framework will evolve as EarthCube grows. Evolving factors are predicted to include, but are not limited to, changing needs of the EarthCube community, impact of emergent technologies, and evolving functions, services, and funding models as EarthCube scales up into a distributed system of systems cyberinfrastructure as NSF envisions it to be.

6.2 IT GOVERNANCE THEORY APPLIED TO EARTHCUBE

A helpful tool to think about EarthCube governance is to map the three questions of effective governance posed by Weill and Ross in *IT Governance* to EarthCube (see Section 5: Status for a review of these questions):

1. ***What decisions must be made?***
2. ***Who should make these decisions?***
3. ***How will these decisions be made and monitored?***

These questions provide us with a useful tool to analyze different components of EarthCube that will likely be addressed by the EarthCube governance, focusing on what decisions must be made to manage these components, who will make these decisions, and how these decisions will be made.

6.2.1 WHAT DECISIONS MUST BE MADE?

First, it is necessary to catalogue the various components of EarthCube. Some of the components of EarthCube identified thus far include science advisory/priorities, technical advisory/priorities, standards development organizations, EarthCube enterprise support, project sponsors portals/cyberinfrastructure, communities of interest/communities of practice, and other stakeholders (long-tail scientists, students and educators, etc.). Figure 11 below describes an initial list of stakeholders, communities of practice, cyberinfrastructure components, existing organizations and government agencies that will likely participate in EarthCube.

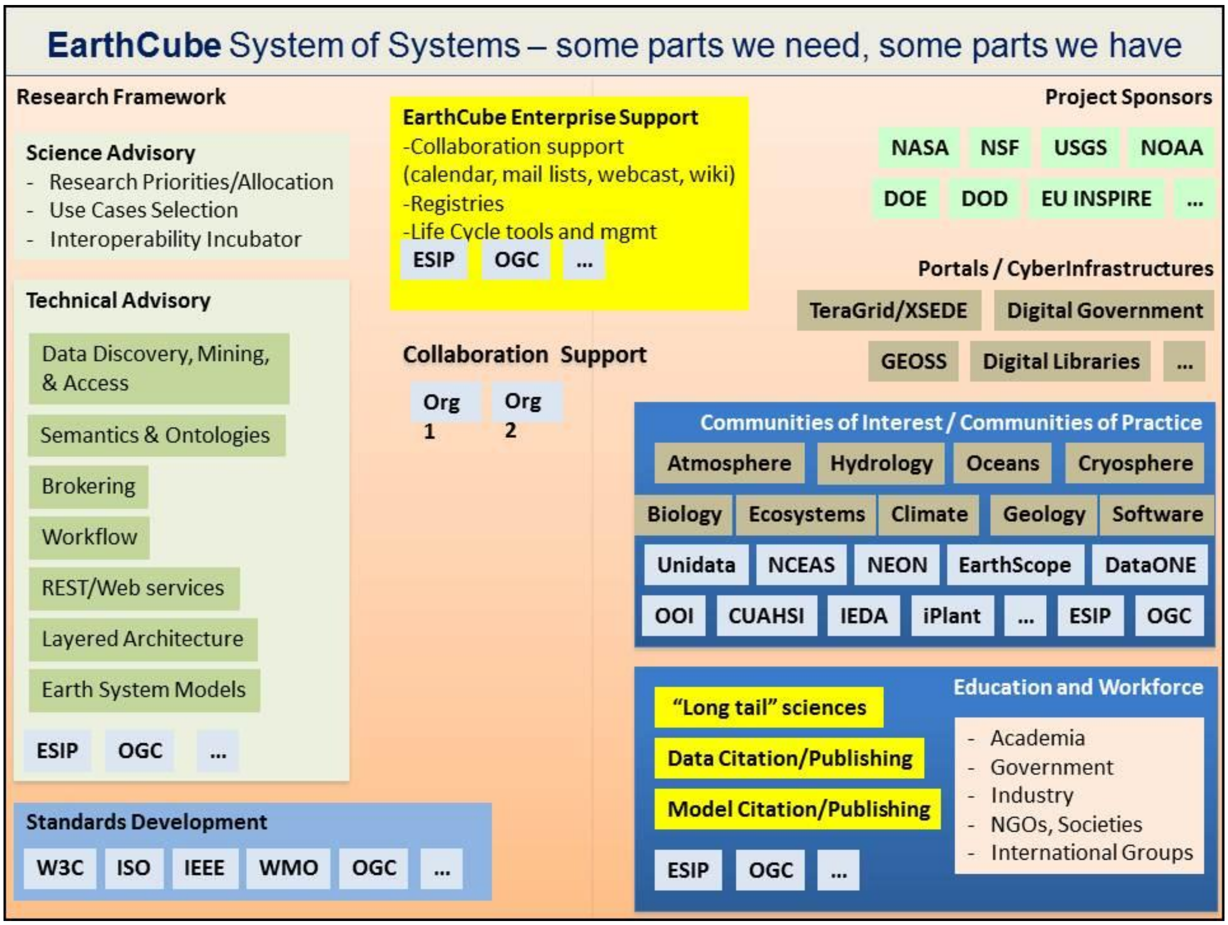

FIGURE 11: A GRAPHICAL REPRESENTATION OF THE MANY COMPONENTS AND COMMUNITIES OF PRACTICE THAT WILL LIKELY COMPOSE EARTHCUBE STAKEHOLDERS.

An EarthCube governance framework must determine how all of these dependent, independent, and interdependent components interact. Important decisions will likely include, but are not limited to, determining leadership roles, setting standards, allocating resources, determining strategic and tactical oversight, coordinating and communicating across EarthCube, and ensuring community needs are met.

6.2.2 WHO MAKES DECISIONS?

The current governance model for EarthCube is one in which NSF funds the initial components (Working Groups and Concept Teams). It has not yet been determined how these components interact with each other. This model is represented in Figure 12 below.

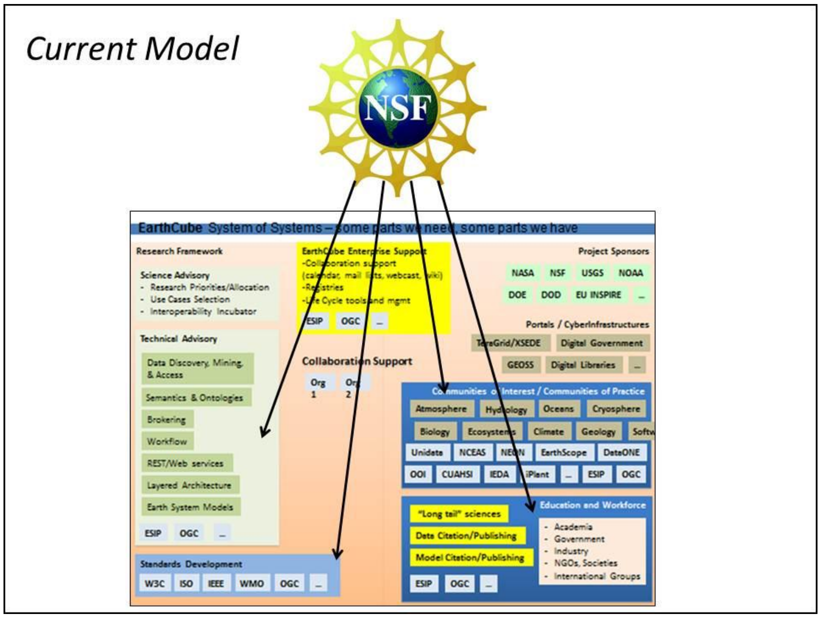

FIGURE 12: CURRENT GOVERNANCE MODEL OF EARTHCUBE.

Figures 13 below represents a centralized governance model as a potential model for EarthCube, in which a centralized office governs distributed EarthCube components.

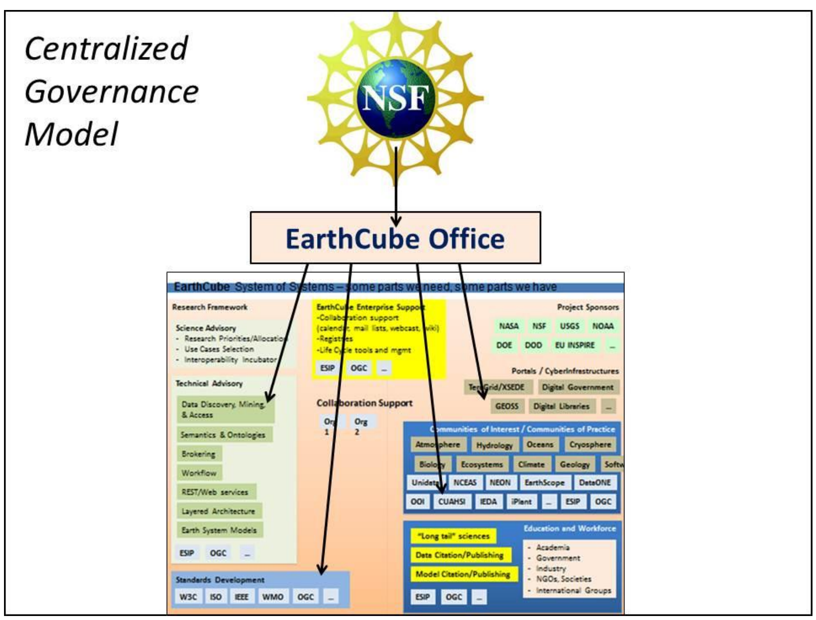

FIGURE 13: A CENTRALIZED MODEL AS A POTENTIAL GOVERNANCE MODEL FOR EARTHCUBE.

Figure 14 demonstrates a distributed governance system as another potential governance model for EarthCube, in which dependent, interdependent, and independent components of EarthCube are not governed or regulated by a central office. This model is analogous to the internet in which there is "…no centralized governance in either technological implementation or policies for access and usage; each constituent network sets its own standards."[198] The internet as an analog to EarthCube is examined in more detail in later in this section.

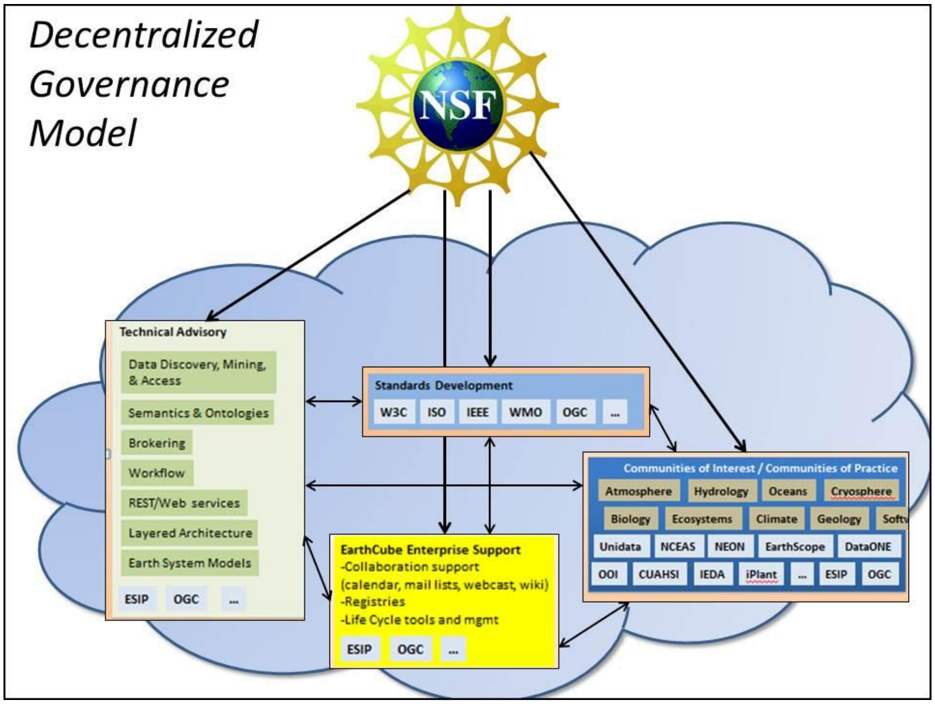

FIGURE 14: A DECENTRALIZED MODEL AS A POTENTIAL GOVERNANCE MODEL FOR EARTHCUBE.

6.1.3 HOW WILL DECISIONS BE MADE AND MONITORED?

The previous section (Section 5: Status) reviewed Weill and Ross' governance matrix, in which different governance models are employed to make different types of decisions within corporate IT governance. Although Weill and Ross focused on governance within individual corporations, it is possible to extrapolate their idea of governance as a system of models to a larger, more complex entity such as EarthCube. In this way, one can envision how different governance models may be employed to manage different EarthCube functions, and how they may come together to form an enterprise-level governance framework.

The EarthCube governance matrix below is a first attempt at examining how the governance models presented in Section 5: Status (top row), may be employed to make different types of decisions within EarthCube components (left column). In this matrix, Weill and Ross' idea of governance a system is broadened in scope to include non IT decisions that may be pertinent to

[198] Wikipedia, "Internet," http://en.wikipedia.org/wiki/Internet, accessed June 1, 2012.

EarthCube, such as organizational principles and goals, organizational investment and prioritization, business model, and science application needs (Table 8). The list of decisions is by no means exhaustive, and will certainly grow as EarthCube governance components, functions and governance needs are identified.

EarthCube Governance Matrix

Governance Model[199] Decision[200]	*Benevolent Dictatorship*	*Domain Science Monarchy*	*IT Monarchy*	*Federal*	*Duopoly*	*Feudal*	*Anarchy*	*Meritocracy*
IT Principles								
IT Architecture								
IT Infrastructure Strategies								
Business Application Needs								
IT Investment and Prioritization								
Organizational Principles and Goals								
Organizational Investment and Prioritization								
Business Model								
Science Application Needs								

TABLE 8: EXPANDED GOVERNANCE MATRIX

Another expanded governance matrix maps possible decision-making processes in atmosphere use case (Table 9).

---

199 Business Monarchy, IT Monarchy, Federal, Duopoly, Feudal, Anarchy archetypes come from Weill and Ross, *IT Governance,* 11; Benevolent Dictatorship, Meritocracy and Monarchy archetypes come from Ross Gardler and Gabriel Hanganu, "Governance Models," Open Source Software Advisory Watch.

200 IT Principles, IT Architecture, IT Infrastructure Strategies, Business Application Needs, and IT Investment and Prioritization come from Weill and Ross, *IT Governance*, 11.

Decision	***IT principles***	***IT architecture***	***IT infrastructure strategies***	***Application needs***	***IT investment***
Archetype					
Benevolent Dictatorship					
Domain Science Monarchy					
IT Monarchy					
Federal			X		X
Community of Practice	X	X		X	
Crowd-source					

TABLE 9: EXAMPLE OF EXPANDED GOVERNANCE MATRIX IN AN ATMOSPHERIC USE CASE.

6.2 EARTHCUBE GOVERNANCE AS A SYSTEM OR SYSTEM OF SYSTEMS

We do not see how any one person or even a small group can comprehend all the aspects of a complex system and undertaking such as complex and diverse as cyberinfrastructure for the geosciences. Thus, in synthesizing research on governance and historical infrastructure development with community feedback on EarthCube, we've come to the conclusion that EarthCube governance must accommodate a system of governance archetypes to meet the needs of each of the EarthCube components and the goals of NSF.

This means EarthCube governance will be comprised of organizational magisteria that may be independent, dependent, or interdependent, as needed to meet the needs of a rapidly evolving system. We anticipate the need for a facile, user-friendly ability to find, access, and comprehend information across the realms of EarthCube and collaborating systems, in order to effectively build, organize, and subsequently manage it as fundamental infrastructure.

At this time we recommend a system of systems approach for EarthCube because:

- Governance has to be flexible to accommodate a system of governance archetypes that meet the needs of each of the sub-discipline communities and achieve the goals of NSF; Governance as a system has the ability to address foreseen and unforeseen challenges
- Governance as a system can also address the requirements sections

Suggestions:

- Take advantage of the testbed environments that have already been developed (OOI, NEON, etc.). A group that identifies approaches that may differ from the existing infrastructure.
- Further engage the long-tail scientists so that they are not driven off.
- Develop metrics based on how active each of the segments of the research community is within EarthCube.,
- Define and accomplish use cases that span domains.

6.3 THE INTERNET AS AN EARTHCUBE ANALOG

An ongoing discussion is whether EarthCube governance is of a single integrated entity or a system of governance of separate, but collective entities. Regardless, EarthCube governance is over a process that is responsible for building infrastructure that becomes ubiquitous and effectively invisible to users. Upon examining historical infrastructure development, we believe the most applicable case studies for EarthCube may thus not be of individual parts of the community, but analogous national infrastructures, like the Internet.

"**The Internet has no centralized governance in either technological implementation or policies for access and usage; each constituent network sets its own standards.**"[201] Critical elements are handled by independent but collaborating organizations – the Internet Corporation for Assigned Names and Numbers (ICANN) directs name spaces and address spaces, while the Internet Engineering Task Force (IETF) maintains the standard core protocols. Other managing bodies of the Internet include (Figure 15):

- ISOC -- Internet Society
- IAB -- Internet Architecture Board
- IRTF -- Internet Research Task Force
- IANA -- Internet Assigned Numbers Authority
- NSI -- Network Solutions
- Accredited Domain Name Registrars

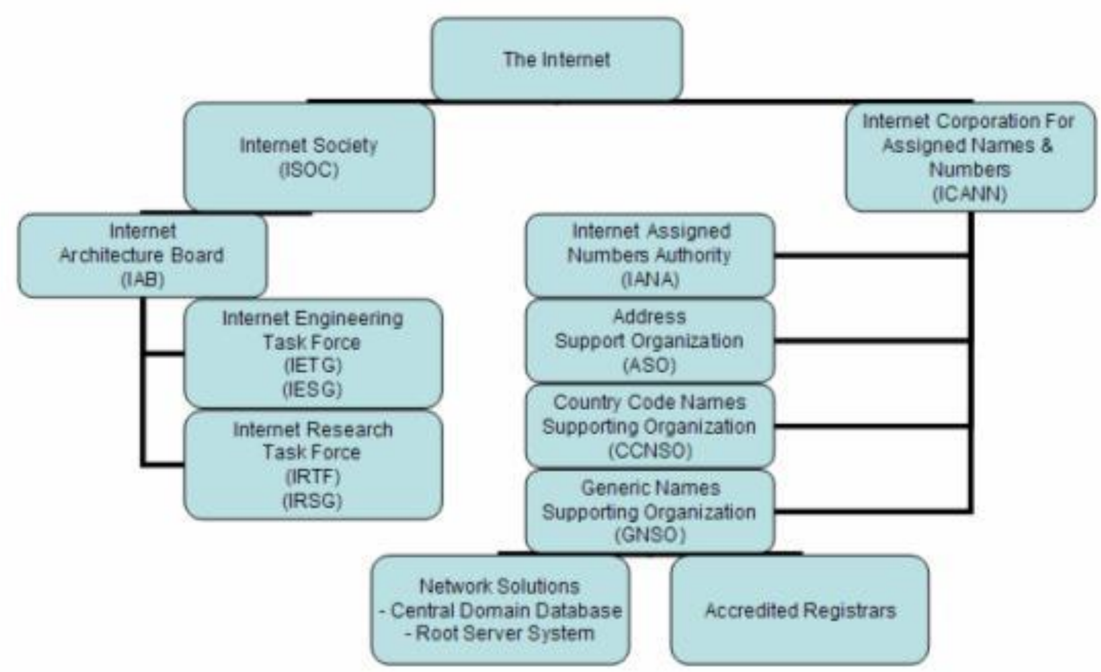

FIGURE 15: INTER-RELATIONS OF SOME OF THE MANAGING BODIES OF THE INTERNET[202]

[201] Wikipedia, "Internet."
[202] Living Internet, "Internet Management," http://www.livinginternet.com/i/iw_mgmt.htm.

The Internet Society notes that **"the Internet itself is comprised of tens of thousands of autonomous networks that voluntarily interconnect by using implementations of these [open Internet] standards**."[203]

"**It is therefore misleading to use the term 'Internet Governance' when 'The Internet' is clearly not a single entity to govern**. It is perhaps more useful and accurate to refer to 'Internet Coordination', as experience has shown that various forms of close coordination are essential to ensure operational stability and preserve architectural integrity. Specifically, close coordination is essential to successfully develop and deploy protocols and carefully allocate the resources that these protocols require to be useful."[204]

Similarly, the Living Internet when asked who manages the Internet, responds that "**It is often said that there is no central control, administration, or management of the Internet**. While this is generally true, there are several well-known organizations that work together in a relatively well structured and roughly democratic environment to collectively participate in the research, development, and management of the Internet…"[205]

---

[203] Internet Society, "Internet Governance: A Misnomer?" 2012, http://www.isoc.org/news/2.shtml, last updated 2012.
[204] Ibid.
[205] Living Internet, "Internet Management," last updated 2012, http://www.livinginternet.com/i/iw_mgmt.htm.

7.0 PROCESS

7.1 GENERAL PROCESS

The timeline described in the Milestones and Tasks (p.12) and Section 8: Timeline (p. 91) has built-in processes to determine the initial EarthCube stakeholder community, identify the governance wants and needs of EarthCube Working Groups and Concept Teams and the greater EarthCube community, identify knowledge gaps in community input and governance research, and identify the initial governance committee to implement initial governance framework.

To continue forward, we recommend implementing Steps 1 and 2 of the Milestones and Tasks:

1. August 31, 2012: Determine the Governance Framework to meet community needs and NSF goals for successful cyberinfrastructure.
2. August 31, 2012: Determine the stakeholder community and identify initial governance committee for engaging the community for input on the governance framework developed in Step 1.

These processes will allow the initial governance committee to develop a draft governance framework based on community needs identified thus far, while concurrently running an aggressive community engagement plan to continually gather governance needs and vet the initial governance framework developed in Step 1. Specific components of developing the governance framework and engaging the stakeholder community are discussed below.

7.2 COMMUNITY ENGAGEMENT

This part of the process will focus on coordination with existing EarthCube Working Groups and Concept Teams, in addition to engaging educators, policymakers, industry, government agencies, the 'long tail' scientists, and other potential EarthCube stakeholders with the goal of identifying their governance requirements for EarthCube. Additionally, discussions on what governance means in general, and how EarthCube may be governed, will be an important component of engaging potential EarthCube stakeholders.

A series of questions have been developed to help EarthCube stakeholders better understand governance and how they likely interact with governance frameworks in their daily life through their work or extracurricular activities:

Governance Ontology [206]

1. How do you manage what you do? Do you have any management bodies (e.g., committees, director, etc.)?
    a. Strategic level
        i. What is their purpose?
            1. Resource allocation, strategic direction, outreach, priorities, fund raising, etc.?
        ii. How many members?
        iii. How often do they meet?
    b. Operational level
        i. What is their purpose?
            1. Design, review, task prioritization, testing, outreach, etc.?
        ii. How many members?
        iii. How often do they meet?
2. Leadership roles
    a. Strategic level
        i. Name and description of roles
        ii. How do you decide how leadership roles are filled?
    b. Operational level
        i. Name and description of roles
        ii. Do you decide how leadership roles are filled?
3. Governance Processes
    a. How do you prioritize tasks?
        i. Who is engaged in prioritization?
        ii. How frequently does it happen?
    b. How do you communicate?
        i. What type?
            1. Telecom, mailing list, wikis, etc.
        ii. Purpose?
        iii. Frequency?
        iv. Is it open?
        v. Mechanisms?
    c. What process do you use to identify requirements?
    d. Dispute adjudication
        i. How do you resolve disputes?
    e. Policies?
        i. Links to policies
    f. Do you have a charter/by-laws/charter?
        i. If so, can it be attached?
4. Are there other aspects of management and organization that should be included that need to be considered for a governance model?

[206] These questions are adapted from Earth System Commodity Governance, http://earthsystemcog.org/projects/cog/governance_object.

Additional questions include:

1. What are the barriers to your participation in EarthCube, if any?
    a. What are your needs and wants in terms of EarthCube governance?
    b. What are your assumptions about EarthCube governance?

7.3 GAP ANALYSIS

Although an extensive background research review was conducted as part of the roadmap writing process, we recognize that through lack of time and resources, there are research gaps that need to be filled.

Areas that require further research include:

1. Social science research on governance
2. Government reports on CI and governance
3. Community Engagement
4. Case Studies on community-building and community-building mechanisms
5. Education and outreach
6. Governance case studies
    a. Analyze and compare governance models of different organizations
    b. Areas of need
    c. Non-profit governance
    d. Governance in Academia
    e. Virtual Governance
    f. Governance in government agencies
    g. Historical case studies
    h. Success stories and failures (lessons learned)

7.4 GOVERNANCE FUNCTIONS

We propose to identify specific functions that EarthCube governance should carry out, without determining how these functions should be carried out, or who is responsible for their implementation. Identification of the EarthCube governance functions will require a synthesis of the Working Group and Concept Team draft roadmaps, community input gathered throughout the roadmap-writing process, and information gleaned from governance research review, in addition to an aggressive community engagement program throughout the second half of 2012.

These functions will form the foundation of the Governance Framework document that will the product of Steps 1 and 2 of the Milestones and Tasks presented in this roadmap. Thus, identification of initial EarthCube governance functions is a high priority following the June 2012 charrette. For example, we know that EarthCube governance should determine leadership roles and define the EarthCube mission, vision and goals. It would be premature at this point to provide prescriptive guidance on how, and who, carries out these functions, however. Much greater knowledge and understanding of EarthCube stakeholder needs is required before these decisions may be made.
Therefore, we recommend the who? and how? of EarthCube governance be addressed as part of Steps 3 and 4 of the Milestones and Tasks (Establish Terms of Reference for EarthCube Governance

and Implement the suggested EarthCube Governance Terms of Reference – Date To-Be-Determined).

7.5 GUIDING PRINCIPLES

In addition to the identification of EarthCube governance functions, we propose to develop a list of Guiding Principles that will guide how EarthCube governance is implemented. Instead of offering specific instructions on how governance should be carried out, these principles will give guidance assure that EarthCube governance is implemented in a manner consistent with community and NSF goals.

8.0 TIMELINE

8.1 TIMELINE RECOMMENDATIONS

We have developed a general timeline for implementing the developing and implementing an initial framework for EarthCube. The primary deliverables are documents reporting the findings and decisions leading to the scope of work for the initial governance framework, a more complete analysis of the community to be served and an aggressive community engagement program to vet the initial governance framework, and an initial terms of reference which can then be incorporated into an NSF solicitation, if NSF chooses to write a solicitation. We envision that Steps 1 and 2 will occur prior to August 31, 2012, although engagement of potential EarthCube stakeholders and vetting of the initial governance framework will occur throughout the second half of 2012. We then foresee an EarthCube governance prototype being implemented in early 2013 that will be responsible for establishing and implementing the initial EarthCube Terms of Reference.

It is the position of this Governance Roadmap Working Group that the roadmap be executed expeditiously and be completed in large part before, not in parallel with, the execution of other roadmaps and as soon as feasible after the second NSF EarthCube Charrette in June, 2012.

8.2 MILESTONES AND TASKS

Goal: To provide a roadmap such that an optimal implementation – agnostic governance framework can be put in place to meet the community needs and achieve the goals of EarthCube.

Step 1: Determine the appropriate Governance Framework to meet community needs and NSF goals for successful cyberinfrastructure

TIME: begin with Charrette, complete by August 31, 2012
DELIVERABLE: Document reporting findings and decisions leading to selection of Initial Governance Framework

1. What Governance models may be best suited for EarthCube?
    a. How will competitive solicitations be enabled, if required?
    b. How will decisions be made?
    c. How will the governance framework be held accountable?
    d. What are the relationships of governance elements to NSF?

2. What are the scopes of authority of governance within and among elements of EarthCube?
    a. Who is in charge of creating and updating EarthCube milestones and determining how they will be achieved? [Key Question for Charrette]
    b. A governance system provides overall steering and oversight for
        i. Leadership
        ii. Adjudication
        iii. Policy development (Science Priorities)
        iv. Intellectual property right protections
        v. Data publishing & data citation

        vi. Model publishing & model citation
        vii. Other use cases
        viii. Sustainability
    c. Spending authority, pass-through funding coordination
    d. Operations
    e. Outreach and communications
    f. Proactive coordination
    g. Architecture maintenance and systems integration support (Technical Priorities)
        i. Standards selection processes supported by:
            1. Prototyping and risk reduction
            2. Testbeds
        ii. Life-cycle data management
            1. Certification of roles
        iii. Security architecture
            1. Authentication
            2. Identity Management
            3. Single sign-on
            4. Threat Mitigation
            5. Policy decision and enforcement
        iv. Registry Management
        v. Other use cases

3. How will the community be involved?
    a. Management and organization - within each community-driven governance element, there could be
        i. Board of Directors
        ii. Advisory Committee structure
        iii. Science Advisory Committee
        iv. Technical Advisory Committee
        v. Program Committees (e.g., Finance, Nominations, Policy, etc.)
        vi. Functional Committees (e.g., Workflow, Data Management, etc.)
        vii. Identification of desired communities of practice. (E.g. Community of practice for outreach and education about current standards, conventions, and best practices for data models, model frameworks, and software engineering. This would be cross-cutting among all science domains, and use social networking tools and outreach.)
    b. End-user engagement & community ownership
        i. Decision and policy makers
        ii. Existing data centers
        iii. Long-tail scientists

4. How will EarthCube governance functions be supported? Any or all of the following methods?
    a. Grants and Contracts Funded
    b. Volunteer
    c. Membership Fees
    d. Public/Private Partnerships

    e. International Initiatives
    f. Provision of Services
5. How will the EarthCube Governance structure manage its own evolution?
    a. What processes are there to recognize new needs as they emerge?
    b. Governance structure should be dynamic

<u>Step 2:</u> Determine the stakeholder community and identify initial governance committee for engaging the community for input on the governance framework developed in Step 1.

TIME: begin with Charrette, complete engagement plan by August 31, 2012, complete initial engagement by beginning of 2013

DELIVERABLE: Document reporting findings and decisions, identifying who is involved

1. Much of this is being done through the current EarthCube community building process that NSF has initiated
    a. Coordination with NSF Concept Awards and Working Groups
    b. Coordination with Focus Groups (e.g. Industry, International, Federal Agencies, Education, etc.)

2. Identify an initial representative group to act for the larger community (EarthCube Governance Committee) to implement the Roadmap
    a. What size should initial group be?
    b. What communities need representation?
    c. What role does NSF play in this?
    d. Determine how this initial group is selected.
    e. Identify linkages to other existing NSF cyberinfrastructure facilities
    f. Leadership -
        i. Who is in charge of making decisions?
        ii. What decisions need to be made?
        iii. Does the individual/group in charge of making decisions change depending on what type of decision is being made?
        iv. Who has the authority to resolve disputes?
        v. What mechanisms are in place to allow for new leadership to emerge if/when necessary?

3. Craft a community engagement program to gather EarthCube governance requirements and vet the initial Governance Framework developed in Step 1
    a. Gather requirements for the earthcube.org website
    b. Engage the geosciences, atmosphere, and ocean communities
    c. Engage the computer science, IT and software communities
    d. Map the EarthCube stakeholder community
    e. Develop materials to market EarthCube to potential new users
    f. Hold webinars for target communities of practice
    g. Engage existing EarthCube Working Groups and Concept Teams
    h. Publish EarthCube articles in trade journals
    i. Engage domain and IT communities through a stakeholder survey
    j. Engage the community through social media

k. Attend conferences of opportunity, such as ESIP and AGU, to engage potential EarthCube stakeholders
l. Employ public listservs, wikis, and the EarthCube Ning site to communicate with EarthCube stakeholders and provide an additional forum for community input

Step 3: Establish Terms of Reference for EarthCube Governance

TIME: Date To-Be-Determined

DELIVERABLE: Initial Terms of Reference for EC Governance Entity that covers the following items

1. EarthCube Governance Committee tasked with drafting an EarthCube Governance Terms of Reference and other documentation that define and codify the Governance characteristics agreed to in the previous steps. This will include but is not limited to:
   a. A process for creating a permanent EarthCube Governance Framework
   b. Decision making structure and leadership
   c. Determination of issues regarding membership, representation, voting, planning
   d. Address issues of evolving EarthCube goals
      i. Who is in charge of creating and updating EarthCube goals? [Key Question for Charrette]
      ii. How does this relate to governance framework?
   e. A transition plan from initial to permanent framework
   f. Participation policy - what does it mean to be a participating EarthCube community member, developer, contributor, etc.?
   g. Process to gather and integrate feedback on proposed EarthCube governance structure from broader community

2. Revise proposed EarthCube Terms of Reference based on feedback

3. Finalize EarthCube Terms of Reference and seek approval of the EarthCube community

Step 4: Implement the suggested EarthCube Governance Terms of Reference

TIME: Date To-Be-Determined

DELIVERABLE: Document reporting decisions; workshop series

8.3 GANTT CHART

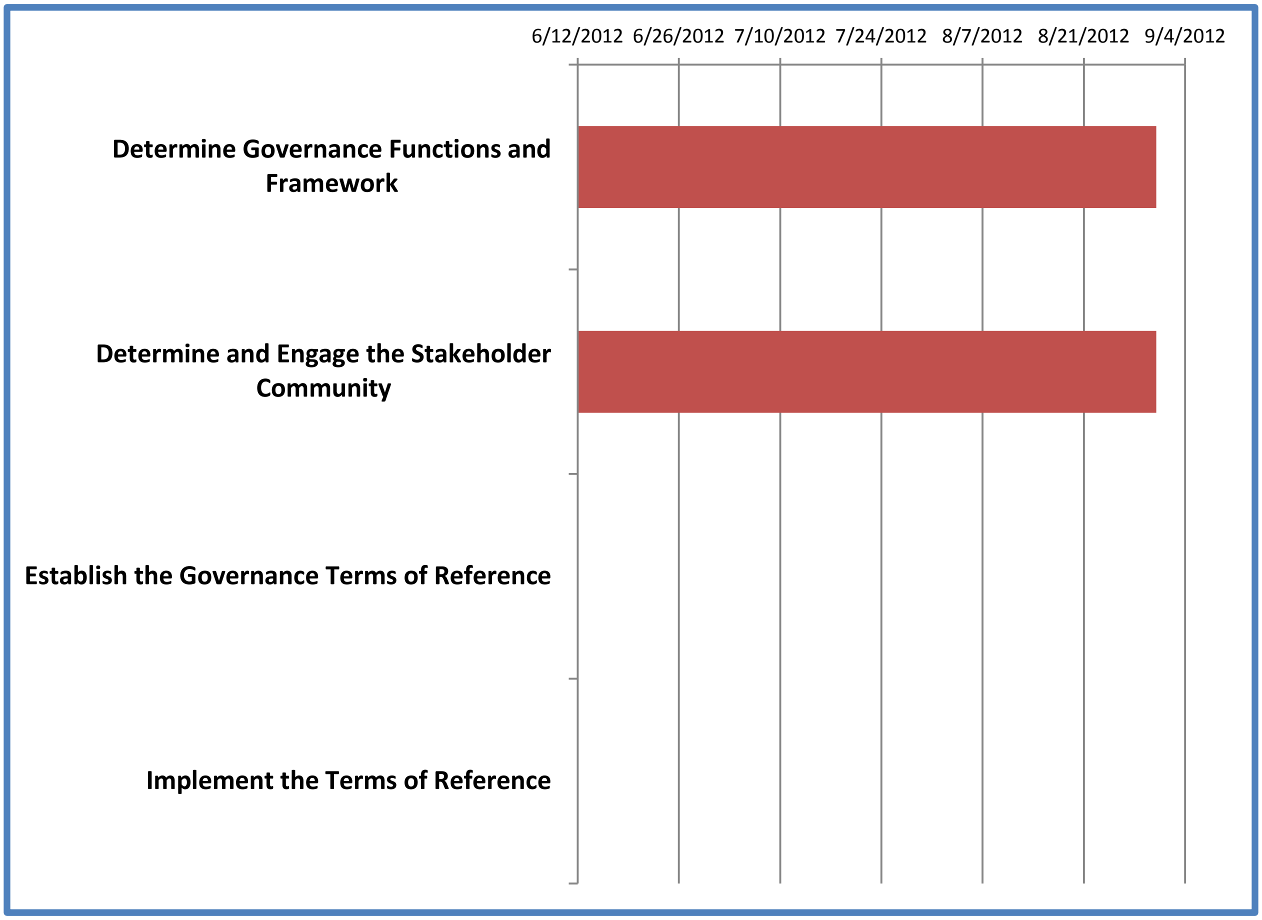

FIGURE 16: GANTT CHART FOR IMPLEMENTATION OF EARTHCUBE GOVERNANCE MILESTONES AND TASKS.

9.0 MANAGEMENT

9.1 MANAGEMENT RECOMMENDATIONS

We suggest that an initial representative group act for the larger community to execute the Governance Roadmap and develop an interim method to establish community concurrence. We suggest that the current Governance Working Group (with membership modifications as noted above) is suitable for this task, and that, upon expanding to include new members, the group be called the Interim Governance Committee (IGC).

We suggest the IGC investigate a system of systems, with distributed responsibilities and authorities, as needed to provide the appropriate mix of competition or alternative approaches, standards development and enforcement.

In the short term, we also feel the IGC needs to identify the current components of cyberinfrastructure (data and service providers), their organizational paradigms, governance needs, the interactions needed among them and between them and systems outside of EarthCube, and the needs of EarthCube consumers (including those comprising the 'long tail' of scientists.

In subsequent discussion, we offer recommendations on the duties and goals of this interim governance body that would be influential on the organization and its scope of work.

9.2 LESSONS LEARNED

In this section we use the term “governance” to encompass the concept of “navigation strategies” offered by Edwards et al in “Understanding Infrastructure,” in which “designing effective navigation strategies will depend on strategic collaborations between social, domain, and information scientists.”[207] They noted that “Transparent, reliable infrastructural services create vast benefits, but there are always losers as well as winners in infrastructure formation. Questions of ownership, management, control, and access are always present.”[208] Although these historical infrastructure dynamics and tensions were presented in Section 5: Status, we believe they deserve additional attention here, focusing on how they related to cyberinfrastructure management.

There are a number of integrative data community initiatives prior and extant that have not gotten the traction that was expected. How do we get synergistic process cutting across the entire community? That social conundrum -- how we will investigate the processes by which existing systems and designs are most effectively integrated, or how we are building the social cyberinfrastructure for collaboration - may be the key to success.

---

[207] Edwards et al., “Understanding Infrastructure,” ii.
[208] Ibid.

A decade ago, there were few attempts to integrate or even collaborate among the myriad cyberinfrastructure projects within any of the geoscience communities, let alone across disciplinary lines. In the past few years, we have seen a transformation in thinking both in funding agencies, and in informatics practitioners. Emphasis has shifted from large centralized single-point-of-access data bases towards data systems hosted by providers in a distributed, web-based network with multiple interfaces accessible to a variety of clients. This has promoted convergence towards open-source, interoperable data networks enabled by emerging information, communication, and computer technology.[209] The Governance Steering Committee members of this project all lead efforts that have adopted these approaches.

Demands for operational cyberinfrastructure come not just from the academic scientific community, but from industry and government for use in socioeconomic and policy decision-making.[210] Mark Abbott in 'A New Path for Science?" notes:

> *"The alignment of science supply and demand in the context of continuing scientific uncertainty will depend on seeking out new relationships, overcoming language and cultural barriers to enable collaboration, and merging models and data to evaluate scenarios…. Capturing the important elements of data preservation, collaboration, provenance, and accountability will require new approaches in the highly distributed, data-intensive research community …[T]his new research enterprise will be more unruly and less predictable, resembling an ecosystem in its approach to knowledge discovery. That is, it will include loose networks of potential services, rapid innovation at the edges, and a much closer partnership between those who create knowledge and those who use it. As with every ecosystem, emergent (and sometimes unpredictable) behavior will be a dominant feature."* [211]

We have seen widespread convergence in recognition of the need to promote networking of geoscience and information technology activities and adherents into a Community of Practice for developing, sharing, using, and maintaining cyberinfrastructure to facilitate data discovery, access, management, utilization, and preservation. The emerging consensus on cyberinfrastructure includes management and organization of software and hardware for implementing web services, specifications for protocols and interchange formats, workflow processes, semantics, ontologies, and brokering, and in addition to the knowledge required to integrate, deploy, operate, and sustain the distributed information system based on these.

"Social and historical analyses reveal some base-level tensions that complicate the work of infrastructural development...including that boundaries between technical and social solutions are mobile, in both directions: the path between the technological and the social is not static and there is no one correct mapping. Robust cyberinfrastructure will develop only when social, organizational, and cultural issues are resolved in tandem with the creation of technology-based services. Sustained and proactive attention to these concerns will be critical to long-term success.[212]

---

209 John Delaney and Roger Barga, "A 2020 Vision for Ocean Science?" in *The Fourth Paradigm*, ed. Tony Hey, Stewart Tansley, and Kristin Tolle (Redmond, Washington: Microsoft Research, 2009), 27-38.

210 Mark Abott, "A New Path for Science?" in *The Fourth Paradigm*, ed. Tony Hey, Stewart Tansley, and Kristin Tolle(Redmond, Washington: Microsoft Research, 2009), 111-116.

211 Ibid., 114-115.

212 Edwards et al., "Understanding Infrastructure, i.

9.3 ROLES AND RESPONSIBILITIES

The goal of this section is to identify potential roles in developing an EarthCube governance framework and engaging potential EarthCube stakeholders to build the EarthCube community and vet the initial framework.

The social cyberinfrastructure needs to lead the technical cyberinfrastructure in order for the technical solutions to have a chance to take hold. Getting the crowd participation model right may be essential to getting the solution right, or agreed to. It is our recommendation that the Governance Roadmap be implemented as early as possible in the EarthCube process, in advance of technical components, rather than in parallel with them. There is an imperative to move promptly before too many "stovepiped" approaches preclude the kind of networking, coordination, and connections that today's cyberinfrastructures can enable right now.

Interim Governance Committee

An initial representative group is needed to act for the larger community to execute the Roadmap and develop an interim method to establish community concurrence. We suggest that the current EarthCube Governance Work Group (with membership modifications as noted above) is suitable for this task.

The questions that the IGC will start with include:

- What are the initial functions that EarthCube should carry out? What process can be used to develop this list so that it reflects the needs of EarthCube stakeholders.
- How should sustainability be approached? What is to be sustained—community facilitation and knowledge development or technological components of the data and information network? Both?
- How can maximum community buy-in be achieved as EarthCube moves forward?
- How can existing resources be leveraged for sustainability? Should a separate organization be formed, or would it be more cost effective to work within an existing organization?
- What kinds of relationships need to be established between geoscience community members, current infrastructure, and the full-range of informatics?
- What is the scope of EarthCube governance?
- What are the artifacts (e.g., specifications, vocabularies, hardware infrastructure, task groups…) that need to be governed? What kind of governance mechanism is needed? How to balance agency/institutional goals and mandates with community needs?

Liaisons to Other EarthCube Groups

Governance is not done in isolation and needs the input of the other EarthCube Working Groups and Concept Teams. The Governance Steering Committee team has identified a set of liaisons to these groups as part of the roadmap-writing process. Throughout this process, the liaisons served as a bridge between the more technically focused groups and the Governance Working Group, and help ensure unique governance issues and needs from each segment of EarthCube were identified and addressed. We will continue to rely on these liaisons for guidance and feedback throughout the development and vetting of the initial Governance Framework document over the next six months.

Community Coordinators

As part of the roadmap-writing process, we identified the role of Community Coordinator, whose responsibility is to dedicate the necessary time and effort required to plan and lay the groundwork to establish a functioning community. Community Coordinator duties include 1) building, supporting, and facilitating professional and social networking paths (working groups, forums, online collaborations, wikis, etc.) among the participants, 2) keeping participants informed of activities, products, developments, and issues emerging among active EarthCube initiatives, and 3) organizing EarthCube governance roadmap activities.

We have identified the need for Community Coordinators for each of the components of NSF's GEO Directorate (Geo, Oceanic, and Atmospheric sciences), the software community, the standards community, the IT community, and the information sciences community. Community Coordinators will play an active role in the community engagement program following the June 2012 Charrette.

EarthCube Stakeholder Community
The targeted EarthCube community is fractal – projects from individual investigators, to teams or groups, to institutions and domain-wide or geographically distributed communities. Each Governance Roadmap Work Group member connects directly from 10s to 100s of others, although they range widely in their cyberinfrastructure engagement or interest. Each of the Governance Steering Committee members is also active in multiple cyberinfrastructure activities, again, providing established conduits to readily and quickly connect with broader communities. The makeup of the Steering Committee covers multiple players from academia, government (both state and federal), and the private sector, domestically and internationally.

We recognize that cyberinfrastructure is being built by individuals, teams, groups, institutions, and communities. Many of these represent or are connected to the projects represented by members on the Governance Steering Committee; however, there are many that are not. The successor to the Governance Steering Committee must keep the governance decision-making and implementation processes open to participation by all interested parties – from individuals to institutions (public and private) to networks, and across the domain, information and computer sciences.

The process to develop the governance framework will continue to be transparent and open to the participation to any and all interested parties. Components of the community engagement plan described in Section 2: Communication will ensure participation of a diverse group of EarthCube stakeholders in the development and revision of the initial EarthCube governance framework.

9.4 SUPPORT

This section documents support processes within the project.

Near Term (Next 6 months)
Cyberinformatics practitioners in the various groups typically interact at irregular intervals, usually by chance, in the context of other meetings. Exchange of information and knowledge between synergistic activities and supportive collaborators is typically ad hoc and anecdotal. One of the goals of the Interim Governance activity is to organize opportunities and events and open communication paths specifically designed to bring different project teams (e.g., technical, scientific, management) together to foster collaboration and leverage effort.

The IGC will support the EarthCube community by: 1) Building, supporting, and facilitating professional and social networking paths (working groups, forums, online collaborations, wikis, etc.) among the participants; 2) Keeping participants informed of activities, products, developments, and issues emerging among active geoinformatics initiatives; 3) and organizing EarthCube governance sessions at existing workshops and conference. Specific tasks will include:

1. Gathering requirements for the earthcube.org website
2. Engaging the geosciences, atmosphere, and ocean communities
3. Engaging the computer science, IT and software communities
4. Mapping the EarthCube stakeholder community
5. Developing materials to market EarthCube to potential new users
6. Holding webinars for target communities of practice
7. Engaging existing EarthCube Working Groups and Concept Teams
8. Publishing EarthCube governance articles in trade journals to keep the EarthCube community up-to-date on the latest developments
9. Engaging domain and IT communities through a stakeholder survey to gather user requirements
10. Engaging the community through social media
11. Attending conferences of opportunity, such as ESIP and AGU, to engage potential EarthCube stakeholders
12. Employing public listservs, wikis, and the EarthCube Ning site to communicate with EarthCube stakeholders and provide an additional forum for community input

Mid-Term (Next 6 Months to 3 Years)
Although it is not determined yet who will carry out these mid-term activities, we feel these activities are important and should be a focus of the interim governance body that will likely be in place in early 2013.

We anticipate the formation of various subcommittees or workgroups to identify nascent issues, address common problems (technical, social, and structural), and develop white papers or best practice documents to recommend solutions to problems identified by the community. Some example topics include metadata profiles, data content models, network protocols, practices for recognition of and respect for intellectual property, data and resource discovery systems, mechanisms to encourage third-party development of web service tools for analyses, sustainable business models for continuing maintenance and evolution of information resources, and integrating the data management life-cycle into the professional and cultural practice of the geosciences and beyond. The community will also develop, document, and implement procedures for stewardship of these 'practice' documents.

By definition, sustainable cyberinfrastructure is an ongoing support service for the entire geoscience and information technology communities and their stakeholders. A goal is to help ensure this comes to fruition sooner, more efficiently, and more effectively, than if left to individual or *ad hoc* interactions in the community. Based in part on conversations with the other EarthCube work groups and concept teams, we forecast these mid-term governance activities:

1. *Leveraging resources*: Each of the Working Groups and Concept Teams in EarthCube is building comparable cyberinfrastructure and community capabilities to the others, along similar architectural and conceptual lines: most are using open source technologies. Most of

us are collaborating on one-to-one bases to seek out best practices, effective tools and procedures, share experiences, and avoid duplication. We should use interim governance body resources to discover and access existing functions, applications, procedures, and standards developed or adopted by others in the cyberinfrastructure communities. We should use the EarthCube interim governance body as a forum for input to new developments to make them more applicable to a wider user audience of geoinformatics practitioners and users.

2. *Facilitating and enhancing emerging technical and structural advances:* A goal for example that is in general common to EarthCube participants is to provide a framework that will enable applications constructed to accept data input using standard web service interfaces to work with any server that provides data using a known interface. Data providers can then publish information by implementing web services without having to simultaneously devote resources to developing client applications to utilize the information. Once a service implementation is in place, addition of content provided by the service is straightforward. Key components of the network are the service specifications, servers and clients implementing the specifications to provide and utilize content, and metadata catalog services that enable users to locate and utilize network resources. The more that the EarthCube members concur on services and specifications, the more readily we can share data, applications, and procedures across and among groups and resources.

3. *Promoting interoperability across the scientific domains:* A goal is to empower and support community development of a web-service based platform and standardized service inventory for interoperable geoscience information services that are easier to maintain and migrate between hosts for long-term stability and availability.

4. *Supporting the promulgation and institutionalization of agreed-upon standards, protocols, and communities of practice incumbent on establishing a permanent and sustainable cyberinfrastructure in the geosciences*: We see increasingly wide-spread common agreements on key components of a shared cyberinfrastructure. These include a catalog system for data discovery; Service definitions that define interfaces for searching catalogs and accessing resources; Shared interchange formats and conventions to encode information in standard ways; Data providers that publish information using standardized services defined by the network; and Client applications enabled to utilize information resources provided by the network.

5. *Support other EarthCube Collaboration areas in defining their own decision making process.* In an effort to provide and facilitate the collaborative nature of EarthCube, the interim governance body should provide support to existing EarthCube collaboration areas; allowing them to help define and structure the governance necessary in accomplishing their specific goals.

9.5 DECISION-MAKING PROCESS

Decision-making typically involves the following steps:

1. Proposal
2. Discussion
3. Vote (if consensus is not reached through discussion)
4. Decision

Role of the Interim Governance Committee

The IGC serves as the decision-making body for implementing the Governance Roadmap. It will choose a chair. Among the duties of the IGC are to determine specific goals, tasks, and priorities to be carried out. This would logically include determining what work groups to form and the most compelling issues to address, but some work groups will likely self-form and ask for IGC logistical support. The IGC will attempt to work by consensus, but can vote on matters as it decides. The IGC may set up subcommittees that include members from outside the IGC and in fact, from outside the EarthCube membership, the latter of which is encouraged.

The IGC will have standing obligations to the project beyond those mentioned above, including increasing diversity, maintaining openness, and establishing links to users in the geosciences.

Any community member can make a proposal for consideration by the community. In order to initiate a discussion about a new idea, they should send an email to the Governance Group list. This will prompt a review and, if necessary, a discussion of the idea. The goal of this review and discussion is to gain approval for the decision or contribution. Since most people in the project community have a shared vision, there is often little need for discussion in order to reach consensus.

Provisions for Flexibility

The EarthCube governance and processes to create it are open to change. The IGC itself will continuously consider adding members or creating additional organizational avenues for participation as EarthCube membership grows. Collaborators from non-geoscience networks will be invited to participate in IGC meetings and discussions as appropriate.

9.6 CONTRIBUTION PROCESS

Feedback gathered during the initial Governance Roadmap process

The Governance Working Group took a multi-faceted approach to solicit and gather input and feedback during the preparation of the roadmap. The Governance Roadmap Steering Committee is composed of the leader or authors of EarthCube Expressions of Interest submitted principally from coalitions to NSF that addressed governance topics, along with representatives from other efforts with a substantive governance component. The collective membership of the steering committee members teams were identified as the EarthCube Governance Forum (open to any interested party) as an extended core community engaged in governance issues. This network of connections provided an initial linkage to a community with partial overlap to the EarthCube community.

The Governance Working Group offered a plenary online workshop at multiple dates and times to better engage a distributed audience. All webinars, calls, slides, notes and documents, including drafts, were posted on the EarthCube Ning site (http://EarthCube.ning.com/) for asynchronous access.

Subsequently, the Governance Working Group identified a formal liaison with each of the other EarthCube work groups and concept teams. The goals were to 1) solicit each groups' insights on their own governance needs that should be considered and incorporated into the Governance Roadmap, and 2) share our own research and conclusions regarding governance for EarthCube. As part of that process, members of the Governance Working Group participated in webinars and conference calls organized by the other groups, discussing our process and the two goals. Following that, a set of questions was sent to the liaisons and Principal Investigators of each work group and concept team, soliciting specific input on governance needs and feedback on the preliminary Governance Roadmap outline.

These steps were directed primarily at the extended EarthCube community. To step out further, the Working Group organized a series of virtual 'breakout' sessions, targeted specifically to international, industry, and government stakeholders. Working Group members and colleagues compiled contact lists for each targeted group, who received an invitation to participate in one of four sessions that occurred over a multi-day period, along with the introductory slides that were shown, and the questions we hoped to answer. Recipients were encouraged to share the invites with their own colleagues. The sessions were nearly identical but we encouraged attendees to participate in the one that would best address their issues during the discussion part of the sessions.

We held two international sessions - one aimed at the Australia-Southeast Asia time zones, and one at the European time zones.

Attendance at the plenary sessions was larger than at the breakouts, but all generated substantive and valuable discussions and feedback. They resulted in new and active participants in subsequent discussions by phone, online, and by email. We captured these interactions in appendices (specifically Appendices 3 and 4) to this report and in a synthesis in the section on Communications.

The Governance Working Group also relied on the EarthCube Ning site to disseminate important information and gain community feedback. Throughout the roadmap-writing process, however, the limitations of the Ning site became more obvious. Postings are organized by time, making it difficult and cumbersome for users to find relevant documents or join discussions. The vast number of items posted and the complexity of the site made it less than optimum. The Working Group delayed pursuit of separate online communication and collaboration tools in deference to NSF's initial plans to move the content on the Ning site to a more robust environment and to keep all the EarthCube materials in one location.

Part of the goals of the Governance Work Group is to analyze the use of social and online media for carrying out our work effectively. This work is underway and will be included in the Work Groups final report with hope it will be valuable for the work of the IGC and other successor EarthCube teams.

Next Steps
Moving forward, we envision multiple facets of information dissemination. First is the use of the NSF EarthCube Ning site as an online community repository for documents, templates, and other print items. Next, we envision the interaction between participants through the implementation of various online collaboration environments such as wikis, e-mail lists, blogs and other social media outlets. In addition, webinars and virtual workshops will be recorded and uploaded to the EarthCube site if possible and online video sites such as Vimeo, where we set up topical playlists gathered from members and related providers.

Engage stakeholder community
We propose that the Interim Governance Committee (IGC) identify critical or priority common challenges around which to assemble voluntary work groups. The work groups will be convened from the broader communities, whether the members are formal members of EarthCube or not. Face-to-face meetings will convene preferably at existing geoscience and IT meetings featuring geoinformatics/cyberinfrastructure or at an EarthCube organized workshop/conference modeled perhaps along the NSF GeoData11 workshop. However the IGC and work groups will meet more frequently via the web through webinars and teleconferences. IGC staff ("Community Coordinator") could provide staff support for the work groups to the extent possible, especially for logistics, but it is expected that each work group will be primarily responsible for preparing their own reports and recommendations; ensuring that each such report is well documented and publicly available.

We envision an aggressive community engagement program to gather EarthCube governance requirements and vet the initial governance framework developed as part of Step 1 of the milestones and tasks presented as part of this roadmap. This plan is described in greater detail in Section 2: Communication.

Encouraging the involvement of additional interested parties
At its very core, EarthCube embodies the involvement of interested parties and investigators at a variety of organizational settings. Members of the IGC include representatives of State and Federal Government, International NGOs, Academia, and Industry. When expanded to include additional project participants and collaborators, the IGC could include Professional Organizations, Collaborative Networks outside of the Geosciences, and the Public. New researchers, post-docs, graduate students and undergraduates will be drawn to the process embodied by EarthCube as it will provide a comprehensive mechanism for research and training. The IGC should actively seek out participation from members of these groups through the work groups and the other EarthCube participants. In many ways, it is the generation of young researchers and students who will provide the vision for EarthCube.

Given the scope and goals of EarthCube, since there are a number of ongoing activities closely related to EarthCube, the governance model will most likely be a governance system (IT Groups, Domain Groups, Federated Communities, Large Facilities – see Section 5: Status). To complete the governance requirements-gathering process discussed in Section 2: Communications and Section 4: Requirements, solicitation to this larger community is needed.

10.0 RISKS

10.1 INTRODUCTION

There is no certainty that EarthCube will succeed as hoped, or that it will do so within the expected time frame and budget. The challenges identified above are extensive and daunting. One only has to look at the past decade to recognize that early efforts to build cyberinfrastructure failed to meet expectations.

As NSF noted in their guidance questions to the work groups, risk is intimately associated with the challenges of implementing EarthCube governance. Rather than repeat those challenges, we attempted to characterize the fundamental risks that would be faced if the challenges are not adequately addressed.

We examine risk in two contexts: 1) That the process of developing roadmaps and the governance roadmap in particular will succeed, and 2) That the goals of EarthCube can be achieved by implementing the roadmap.

10.2 RISKS TO THE ROADMAP PROCESS

NSF Process

The new approach is a revolutionary departure from the traditional NSF approach. NSF has encouraged a community-driven approach to EarthCube thus far, relying on the community to define and develop EarthCube components. A risk is that the community is not ready to do so.

Group Think and Perceptions of the Committee

One obvious risk for developing the governance structure is to avoid the tendency for group think. This happens when the desire for harmony in the team overrides a realistic appraisal of alternatives. As our team developing the governance roadmap has found out, it is often easy for a dominant idea or proponent to become the standard model and alternatives dismissed as not the best idea without full and fair consideration.

By nature the roadmap-writing process relied on consensus-driven decisions and because of group think and time constraints, there is a tendency for creating an incomplete survey of alternative ideas. There has not been time to examine risks in detail of any of the preferred considered choices.

Inability to Decide Upon Scope and Structure of Governance Process

A risk is that the community will not be able to come together to determine the scope and structure of the EarthCube governance framework.

Short Term Thinking

A risk is selecting a governance framework for the short-term that does not enable EarthCube to evolve as it grows throughout the long-term.

10.3 RISKS IN MEETING EARTHCUBE GOALS

The governance structure needs to be based on the objectives of EarthCube. Unfortunately the objectives of EarthCube are not yet well defined. The November 2011 NSF Charrette tried to outline the capabilities of EarthCube but no attempt was made to prioritize the capability requests and develop clear objectives during the life of the EarthCube project. The highest priority of all of the EarthCube roadmap working groups will be to outline their most important capabilities so a preliminary set of objectives can be created (Figure 17).

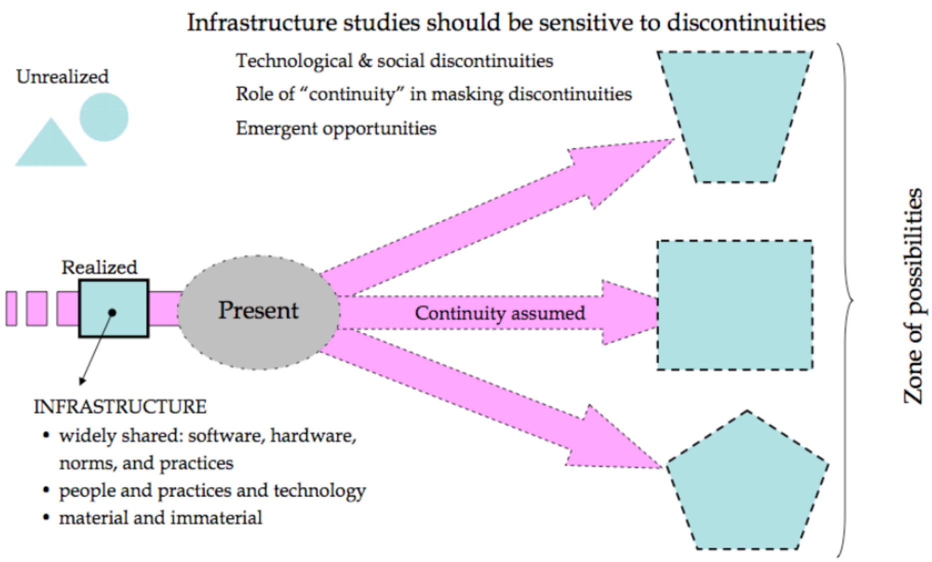

FIGURE 17: VISUALIZING PATH DEPENDENCE AND DISCONTINUITY (GRAPHIC PREPARED BY TROND JACOBSEN IN EDWARDS ET AL, 2007, P 18

Lack of Lessons Learned

One of the most significant risks is using the literature for evaluating best governance models for data infrastructure. Very few "Best Practice Documents" outline the problems and failures of a specific approach. Everyone wants to talk of the success not the project failure. Even published projects that were not totally successful are written to sound like they met all project goals. This is not dissimilar to the political spin doctors that rewrite history to their political group's benefit.

Contingency Plans

A cursory review of the different NSF roadmap groups suggest that no one has started to plan for contingency plans for partial or total failure of the recommended procedures. None of the EarthCube roadmaps to be developed at the Charrette will be entirely complete or even have the best recommendations. For example, the Governance roadmap team has put a timeline together and it or a new governance team may not to be able to meet the goals. Since the entire EarthCube program depends on the governance structure, what contingency plans need to be in place if and when goals are not met on time?

Interweaving Architectures
There is a risk of interweaving the different architectures of different domains (OOI, Earthscope, DataONE, etc.). Each domain has its own architecture; there are numerous technical and social challenges to interweaving different architectures.

Tyranny of the Large Over the Small
A significant challenge and risk for EarthCube is the tyranny of the big projects (large funding) over the smaller funded projects. There is a high probability that the large funded projects will dominate the governance structure. The risk is that the smaller funded project's (more common NSF projects) needs will be ignored and the EarthCube project could fail.

Governance Approach Proves Unwieldy or Ineffective
A risk is that the governance framework of EarthCube will prove to be unwieldy or ineffective.

Sustainability and Evolution of Governance
Early published NSF EarthCube documentation in 2011 discussed the need for sustainability of EarthCube. A predominantly NSF funded EarthCube model would probably have a different governance structure than one that needed financial support from members, industry, foundations, and other government agencies. The governance structure must understand the risks of changing the funding sources.

10.4 RISKS IN INSTITUTIONAL COMMITMENT AND SUPPORT

Development of sustainable infrastructure is a multi-year to multi-decade process, so there is risk associated with NSF's commitment to EarthCube and the ability to maintain the funding stream to bring it to fruition. EarthCube is drawing widespread support within NSF including at the Director's level, and outside with the White House and others, but we realize that people retire, or move, that funding from Congress is not guaranteed, or that priorities change at the institutional and national levels

10.5 RISKS IN COMMUNITY COMMITMENT

The community is solidly engaged in EarthCube as evidenced by participation of 1,000 individuals, the huge amounts of volunteer time being dedicated to the roadmap process, and the enthusiasm at which the Working Groups and Concept Teams formed and dove into interlocking collaborative tasks.

The risk is that the momentum may be difficult to maintain let alone expand to larger communities, due to burnout, over-commitment, or lack of satisfying results.

10.6 RISK MITIGATION

Roadmap process

We have attempted to mitigate the identified risks in part by aggressively engaging the extended governance interest community, the wider EarthCube community, and targeting key focus groups outside of NSF's core constituencies. We have also gone to great lengths to make the entire roadmap process completely open and transparent. While these conceptually are appropriate, it is clear that we still reached only a limited part of the stakeholder community.

Achieving EarthCube goals

The debate over a central governing body vs. a decentralized system has potentially serious consequences. Proponents of centralized governance cannot envision success any other way. Proponents of a decentralized system point to all other successful infrastructures as evidence and express doubt about their participation in a centrally managed organization.

One useful best mitigation strategy to achieving EarthCube goals is to think about prototyping and evaluation of case studies. Use existing prototyping architectures to take advantage of the best available models to address the problem at choosing the best (i.e., most appropriate) models. Each community has its own competing approach. Social science can come in and figure out how to mediate.

Institutional Commitment and Support

Although EarthCube currently depends on NSF as its principle funding source, as EarthCube expands and the stakeholder community grows, it will likely be necessary to seek additional funding sources outside of NSF.

Community Engagement

An aggressive community engagement program will need to take place to ensure that EarthCube will meet community needs, and to promote community buy-in to EarthCube to the greatest extent possible.

APPENDICES

APPENDIX 1: GOVERNANCE RESEARCH REVIEW

EXECUTIVE SUMMARY

This paper acts as an appendix to the EarthCube Governance Steering Committee's Roadmap for Achieving EarthCube Governance by providing a comprehensive review of current governance theories and practices from the sciences, information technology sector, and social sciences. We provide a brief introduction on governance, its importance, and then provide a number of concepts and definitions. Lessons learned from historical infrastructure case studies and governance case studies provide an overview of the challenges faced by a lack of governance or unclear governance structure. Finally, we review recommendations from the geoinformatics and EarthCube communities.

The research review is intended to provide a summary of best practices, concepts, and mechanisms for governance in order to provide the foundation for the next stages of development in EarthCube governance. The most important findings are summarized below.

Section 1: Introduction to Governance. Reviews governance definitions and concepts; compares and contrasts corporate IT governance, non-profit governance, and open source software governance. **Governance definitions include**:

1. IT governance "specif[ies] the decision rights and accountability framework to encourage desirable behavior in using IT. IT governance is not about making specific IT decisions–management does that–but rather determines who systematically makes and contributes to those decisions."[213]
2. Governance "provid[es] the structure for determining organizational objectives and monitoring performance to ensure that objectives are attained."[214]
3. A "governance model describes the roles that project participants can take on and the process for decision making within the project. In addition, it describes the ground rules for participation in the project and the processes for communicating and sharing within the project team and community."[215]
4. Governance "aligns an organization's practices and procedures with its goals, purposes, and values. Definitions vary, but in general governance involves overseeing, steering, and articulating organizational norms and processes (as opposed to managerial activities such as detailed planning and allocation of effort). Styles of governance range from authoritarian to communalist to anarchical, each with advantages and drawbacks."[216]

Effective governance must address three key questions:

Question 1: What decisions must be made to ensure effective management and use of IT? In corporate IT governance, five types of decisions were identified:

1. IT principles: Clarifying the business role of IT
2. IT architecture: Defining integration and standardization
3. IT infrastructure: Determining shared and enabling services
4. Business application needs: Specifying business need for purchased or internally developed IT applications

[213] Peter Weill and Jeanne Ross, *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results, (Boston, MA: 2004)*, 2.
[214] Organization for Economic Cooperation and Development, Directorate for Financial, Fiscal and Enterprise Affairs, *OECD Principles of Corporate Governance*, (April 1999), 5, 219, quoted in Weill and Ross, *IT Governance*, 4-5
[215] Ross Gardler and Gabriel Hanganu, "Governance Models," Open Source Software Watch, last modified February 14, 2012, http://www.oss-watch.ac.uk/resources/governanceModels.xml.
[216] "Governance," EarthSystem Commodity Governance Project, last modified 2012, http://earthsystemcog.org/projects/cog/governance_object

5. IT Investment and prioritization: Choosing which initiatives to fund and how much to spend

Non-profit, open source software and other types of governance have a different set of decisions that must be made.

Question 2: Who should make these decisions? Governance archetypes decision-making rights and authority to individuals or groups of people depending on the type of archetype employed:

1. Business Monarchy: top managers
2. IT Monarchy: IT specialists
3. Feudal: each unit makes independent decisions
4. Federal: combination of corporate center and business units with or without IT people involved
5. IT Duopoly: IT group and one other group (ex: top management or business unit leaders)
6. Anarchy: isolated individual or small group decision making
7. Meritocracy: distributed control is awarded in recognition of contributions to the project
8. Technical Utopianism: technical approach to information management.

Question 3: How will these decisions be made and monitored? The governance archetype best suited for each type of decision is selected. For example, a Business Monarchy might be best for making IT Principles decisions, but a Duopoly might be most appropriate for IT Investment decisions. The matrix below (Table 11) is a tool to map the governance archetype (left column) to the type of decision being made (top row).

Decision Archetype	***IT Principles***	***IT Architecture***	***IT Infrastructure Strategies***	***Application Needs***	***IT Investment***
Business Monarchy					
IT Monarchy					
Feudal					
Federal					
Duopoly					
Anarchy					
Technocratic Utopianism					
Meritocracy					

TABLE 11: "GOVERNANCE AS A SYSTEM ARRANGEMENTS MATRIX—WHICH GOVERNANCE ARCHETYPES ARE USED FOR DIFFERENT TYPES OF DECISIONS?"[217]

Governance mechanisms implement the governance framework on a day-to-day basis. Mechanisms identified in the research review include:

1. **Decision-Making Structures:** Organizational units and roles responsible for making IT decisions, including committees, executive teams, and business/IT relationship managers.
2. **Alignment Processes:** Formal processes that ensure daily behaviors are consistent with IT policies, and provide input for IT decisions. They include IT investment proposal and evaluation processes, architecture exception processes, service-level agreements, chargeback, and metrics.
3. **Communication Processes:** Announcements, advocacy, channels, and education efforts that inform the organization of IT governance principles and policies, and outcomes of IT decision-making processes.

[217] Modified from Weill and Ross *IT Governance*, 2004, MIT Sloan School Center for Information Systems Research (CISR), 2003. Cited with permission in Weill and Ross, *IT Governance*, 11.

4. **Community-Building Processes:** Processes to build virtual and face-to-face communities.

Characteristics of top corporate governance performers :

1. More managers in leadership positions are able to describe IT governance
2. Senior management is engaged in governance and communicates to the rest of the organization
3. More direct involvement of senior leaders in IT governance
4. Clear business objectives for IT investment
5. Differentiated business strategies
6. Fewer renegade decisions and more formally approved exceptions
7. Fewer changes in governance from year to year

Characteristics of top not-for-profit governance performers

1. Executive committees focused on all key assets including IT
2. IT council comprising business and IT executives
3. IT leadership committee comprising IT executives
4. Architecture committee
5. Tracking of IT projects and resources consumed
6. Business/IT relationship managers

Why is Governance Important?

Governance can help assure that decisions throughout the enterprise are consistent with the organizational direction and can, in fact, influence how that direction is established. Effective governance aligns daily behaviors, decision-making processes, and communication mechanisms to meet overall organizational goals.

Corporate IT Governance:[218]

1. **Good IT governance pays off because IT is expensive.** Focused IT spending on strategic priorities can reduce IT costs by involving the right people in the IT decision-making process.
2. **IT is pervasive.** "Well designed IT governance arrangements distribute IT decision making to those responsible for outcomes."[219]
3. **New information technologies bombard enterprises with new business opportunities.** Governance can also take advantage of opportunities created by new technologies by having the foresight to establish the right infrastructure at the right time to create new business opportunities.
4. **IT Governance is critical to organizational learning about IT value.** Effective governance creates mechanisms to debate the potential value of specific IT initiatives and formalize learning about IT and thus is critical for the organization to learn about IT value.
5. **IT Value depends on more than good technology.** "Successful firms involve the right people in the process. Having the right people involved in IT decision making yields both more strategic applications and greater buy-in. These more involved people then produce better implementations."[220]
6. **Senior management has limited bandwidth.** "Decisions throughout the enterprise should be consistent with the direction in which senior management is taking the organization. Carefully designed IT governance provides a clear, transparent IT decision-making process that leads to consistent behavior linked back to the senior management vision while empowering everyone's creativity."
7. **Leading enterprises govern IT differently.** "Top performing firms balancing multiple performance goals had governance models that blended centralized and decentralized decision making. All top performers'

[218] Weill and Ross 2004, *IT Governance:* 14-18.
[219] Ibid., 15
[220] Ibid., 17

governance had one aspect in common. Their governance made transparent the tensions around IT decisions such as standardization versus innovation."[221]

Information Management

Governance also improves information management.[222] Information politics play an increasingly important role in modern for-profit and not-for-profit organizations, universities, virtual organizations, and cyberinfrastructure (CI) projects. As jobs become defined by the information they possess, people may be less likely to share that information. A governance framework allows firms, organizations, universities and other entities to address not only the information itself, but the organizational structure that leads to the generation of that information, and the processes that determine how that information is handled.[223]

Open Source Governance

Effective governance in open source software projects also has a series of benefits. An effective governance framework increases the agility of the project, enhances project sustainability, helps maintain a clear sense of direction, allows control to remain where the leaders want it to be, and encourages potential contributions and project grow.[224]

Section 2: Historical Infrastructure Development Case Studies and Lessons Learned

Reviews common steps, dynamics and tensions in historical infrastructure development[225], and historical case studies: Project Mohole,[226] and leadership principles in the National Science Foundation Network (NSFNET), Unidata, Digital Library for Earth System Science (DLESE), and the National Digital Science Library (NSDL).[227]

Lessons Learned from Historical Infrastructure

1. **Common steps in historical infrastructure development:** 1) system building, 2) technology transfer, 3) integration and consolidation
2. **Historical infrastructure dynamics:** Reverse salients, gateways, path dependence, scale effects
3. **Historical infrastructure tensions:** 1) Time, Scale, and Agency, 2) Winners and Losers, 3) Interest and Exclusion, 4) Ownership and Investment, 5) Data Cultures and Data Tensions, 6) Data Storage and Meaning, 7) Data Sharing, 8) Trust.

Cyberinfrastructure design recommendations:

1. **Purpose.** CI design should permit collaborative work that enables scientists from different domains to work together.
2. **Boundary Work.** CI can bridge the 'great divide' between system building and social analysis.
3. **Principles of Navigation.** It is necessary to establish a set of design principles of navigation instead of a rigid roadmap because there is no one way to design CI.
4. **Standards and Flexibility.** Flexibility should be included in CI design to avoid path dependence and to facilitate change.

Recommendations for Future CI Projects

---

[221] Ibid., 18.
[222] Thomas Davenport, Robert Eccles, and Laurence Prusak, "Information Politics," 1992.
[223] Ibid.
[224] Ross Gardler and Gabriel Hanganu, "Governance Models."
[225] Paul Edwards, Steven Jackson, Geoffrey Bowker, and Cory Knobel, "Understanding Infrastructure: Dynamics, Tensions, and Design - Report of a Workshop on "History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures," 2007.
[226] See D.S. Greenberg, " Mohole: the Project That Went Awry," *Science*, 143, no. 3602, (1964): 115-119,
[227] David Fulker, "Collaboration, Alignment and Leadership, invited piece in 'NDSL Reflections," an online series accessed at http://nsdlreflections.wordpress.com, 2008.

1. **Learn from successes and failures of past CI projects.** A comparative social scientific analysis of CI projects, use of accurate and realistic reporting mechanisms, aligning extant or creating new incentive structures, and 'instrumenting' CI projects for social scientific analysis will promote learning from past CI projects.
2. **Improve current CI analysis.** Infrastructural diagnostics, training for information managers, cross-disciplinary symposia for early-career scientists, and fitting funding to CI-development time scales can improve current CI analysis.
3. **Enhance CI resiliency and research.** CI resiliency, sustainability and reach can be enhanced through design flexibility, mechanisms to incorporate under-represented groups and institutions, and alliances with niche organizations to eliminate redundancies in expertise.

Historical Cast Studies: The failure of Project Mohole can be attributed to a lack of governance framework to clearly define leadership goals, assign leadership authority and accountability, and processes for dealing with disputes and other problems that arose. Success of NSFNE and Unidata is attributed to adherence to clear leadership principles (disciplined thought and action, persistent core, and primary alignment to operational excellence, customer relations, or product leaders , while DLESE and NSDL did not succeed due to lack of clear leadership.

Section 3: Current State of Governance

Reviews case studies in domain science, IT, federated and large Facilities groups, and governance recommendations from recent geoinformatics reports. Case studies demonstrate that governance is as varied as the institutions it serves, even within the same communities. Case studies include:

1. **Domain science**: CUAHSI, IRIS, Unidata
2. **IT:** Apache Software Foundation, DataNet Federation Consortium (iRODS data grid), DataONE, W3C Consortium
3. **Federations:** ESIP, OGC
4. **Large Facilities:** GENI – Global Environment for Network Innovations, SURA US IOOS Coastal and Ocean Modeling Testbed

Geoinformatics reports recommend a balance of leadership and governance to successfully advance the geoinformatics agenda and to highlight four key geoinformatics questions: 1) I can't integrate what I can't find, 2) I can't use something I don't understand, 3) I don't want to use something I don't trust and 4) I can't use something that isn't there anymore.

Section 4: EarthCube Governance Recommendations from Research Review

Reviews thematic and topical governance recommendations from EarthCube White Papers (Governance and Design categories) and outside sources.

Thematic Recommendations:

1. Clarity
2. Collaboration
3. Commodity Governance and Communities of Communities
4. Community
5. Communities of Interest
6. Communities of Practice: Informaticists, Data Generators, Synthesizers
7. Community-based governance
8. Diverse Representation in Governance Framework
9. Evolution
10. Form Should Follow Function
11. Goals
12. Leadership
13. Open Source Software Governance and Open Science
14. Public-Private Partnership
15. Success

16. Sustainability
17. Trust

Topical and Case Study Recommendations

1. AGORA-net
2. Centralized Database
3. CZO Data.
4. CUAHSI-HS
5. DataONE
6. EarthCube as a Double Loop Organization (Virtual Democracy Blog recommendation)
7. EC Network
8. Federation of Earth Science Information Partners (ESIP)
9. ESIP Discovery Cluster
10. Incorporated Research Institutions for Seismology (IRIS)
11. Open Geospatial Consortium (OGC)
12. Professional Networking Service
13. QuakeSim
14. Scientific Forecasts
15. Space-Time Location
16. UNAVCO
17. Unidata
18. USGIN
19. USGS
20. WInSAR Consortium

SECTION 1: INTRODUCTION TO GOVERNANCE

This section answers two key questions: *What is governance? Why is it important?* The section examines key governance definitions, concepts, components, and examples of governance. It then presents characteristics of top governance performers, and compares the governance framework of the best and worst governance performers. The section concludes with a summary of reasons why governance is important.

DEFINITIONS OF GOVERNANCE

In the Governance Roadmap we define ***governance*** as the processes, structure and organizational elements that determine, within an organization, how power is exercised, how stakeholders have their say, how decisions are made, and how decision-makers are held accountable. This definition is an amalgamation of the various background research sources with specific guidance from community input.

The foundations of this definition are discussed here, as we review governance definitions, concepts, models/archetypes, and mechanisms in the context of corporate and IT governance, non-profit governance, and open source software governance. Literature reviewed includes *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results* (Weill and Ross, 2004), "Information Politics" (Davenport, Eccles and Prusak, 1992), "Governance Models" (Open Source Software Advisory Watch, 2012).

Although each source defines governance in a slightly different way, each definition contains elements that are relevant to EarthCube. Several governance definitions are stated below:

1. IT governance "specif[ies] the decision rights and accountability framework to encourage desirable behavior in using IT. IT governance is not about making specific IT decisions−management does that−but rather determines who systematically makes and contributes to those decisions."[228]
2. Governance "provid[es] the structure for determining organizational objectives and monitoring performance to ensure that objectives are attained."[229]
3. A "governance model describes the roles that project participants can take on and the process for decision making within the project. In addition, it describes the ground rules for participation in the project and the processes for communicating and sharing within the project team and community."[230]
4. Governance "aligns an organization's practices and procedures with its goals, purposes, and values. Definitions vary, but in general governance involves overseeing, steering, and articulating organizational norms and processes (as opposed to managerial activities such as detailed planning and allocation of effort). Styles of governance range from authoritarian to communalist to anarchical, each with advantages and drawbacks."[231]

Weill and Ross make a distinction between behavioral (interactions) and normative (rules) sides of governance. They refer to Stilpon Nestor (Head of the Corporate Affairs Division of the Organization for Economic Development), who defines behavioral governance as "the relationships and ensuing patterns of behavior between different agents in a limited liability corporation; the way managers and shareholders but also employees,

[228] Weill and Ross, *IT Governance*, 2
[229] Organization for Economic Cooperation and Development, *OECD Principles of Corporate Governance* quoted in Weill and Ross, *IT Governance*, 4-5
[230] Ross Gardler and Gabriel Hanganu, "Governance Models,"
[231] "Governance," EarthSystem Commodity Governance Project, last modified 2012, http://earthsystemcog.org/projects/cog/governance_object

creditors, key customers, and communities interact with each other to form the strategy of the company."[232] Weill and Ross argue that behavioral IT governance refers to formal and informal relationships as well as assigning decision rights to specific individuals or groups of individuals.

Nestor defines normative governance as "Refer[ring] to the set of rules that frame these relationships and private behaviors, thus shaping corporation strategy formation. These can be the company law, securities regulation, listing requirements. But they may also be private, self-regulation."[233] Weill and Ross argue that normative IT governance encompasses mechanisms for formalizing relationships as well as rules and operating procedures that ensure organization objectives are met.

GOVERNANCE COMPONENTS

A governance framework is composed of a series of decision-making models employed for different types of decisions that need to be made. The following section is taken from Weill and Ross' IT Governance and Davenport et al.'s Information Politics. Although these sources focus on corporate IT governance, the governance archetypes and types of decisions that need to be made can be applied to any entity that relies on IT. Specific examples of how corporate IT governance is described below in 'Corporate IT Governance (see page 15).

Effective governance addresses three key questions:[234]

1. What decisions must be made to ensure effective management and use of IT?
2. Who should make these decisions?
3. How will these decisions be made and monitored?

GOVERNANCE DECISIONS

Governance decisions answers question 1: *What decisions must be made to ensure effective management and use of IT?* Weill and Ross identify five key types of decisions that must be made:[235]

6. **IT principles**: Clarifying the business role of IT
7. **IT architecture**: Defining integration and standardization
8. **IT infrastructure**: Determining shared and enabling services
9. **Business application needs**: Specifying business need for purchased or internally developed IT applications
10. **IT Investment and prioritization**: Choosing which initiatives to fund and how much to spend

These decisions are described in detail below:[236]

- IT Principles
  - What is the enterprise's operating model?
  - What is the role of IT in the business?
  - What are IT-desirable behaviors?
  - How will IT be funded?
- IT architecture
  - What are the core business processes of the enterprise?
  - How are they related?

[232] Stilpon Nestor, "International Efforts to Improve Corporate Governance: Why and How," *Organization for Economic Cooperation and Development*, 2001, 1
[233] Stilpon Nestor, "International Efforts to Improve Corporate Governance: Why and How," *Organization for Economic Cooperation and Development*, 2001, 1
[234] Weill and Ross, *IT Governance*, 10
[235] Ibid., 10-11.
[236] Ibid., 54-55.

    - What information drives these core processes? How must the data be integrated?
    - What technical capabilities should be standardized enterprise-wide to support IT efficiencies and facilitate process standardization and integration?
    - What activities must be standardized enterprise-wide to support data integration?
    - What technology choices will guide the enterprise's approach to IT initiatives?
- It infrastructure
    - What infrastructure services are most critical to achieving the enterprise's strategic objectives?
    - For each capability cluster, what infrastructure services should be implemented enterprise-wide and what are the service-level requirements of those services?
    - How should infrastructure services be priced?
    - What is the plan for keeping underlying technologies up to date?
    - What infrastructure services should be outsourced?
- Business Application needs
    - What are the market and business process opportunities for new business applications?
    - How are experiments designed to assess whether they are successful?
    - How can business needs be addressed within architectural standards? When does a business need justify an exception to a standard?
    - Who will own the outcomes of each project and institute organizational changes to ensure the value?
- IT Investment and prioritization
    - What process changes or enhancements are strategically most important to the enterprise?
    - What are the distributions in the current and proposed IT portfolios? Are these portfolios consistent with the enterprise's strategic objectives?
    - What is the relative important of enterprise-wide versus business unit investments? Do actual investment practices reflect their relative importance?

Governance Archetypes

Governance archetypes answer the question: Who should make these decisions? Weill and Ross identify six governance archetypes that determine who makes decisions:[237]

1. **Business Monarchy**: top managers
2. **IT Monarchy**: IT specialists
3. **Feudal**: each unit makes independent decisions
4. **Federal**: combination of corporate center and business units with or without IT people involved
5. **IT Duopoly**: IT group and one other group (ex: top management or business unit leaders)
6. **Anarchy**: isolated individual or small group decision making

These governance archetypes are explained in detail in Table 12.

IT Governance Archetypes

STYLE	WHO HAS DECISION OR INPUT RIGHTS?
Business Monarchy	A group of business executives or individual executives (CxOs [C-level executives]). Includes committees of senior business executives (may include CIO). Excludes IT executives acting independently
IT Monarchy	Individuals or groups of IT executives
Feudal	Business unit leaders, key process owners or their delegates

[237] Weill and Ross, *IT Governance*, 10

Federal	C-level executives and business groups (e.g., business units or processes); may also include IT executives as additional participants. Equivalent of the central and state governments working together
IT Duopoly	IT executives and one other group (e.g., CxO or business unit or process leaders
Anarchy	Each individual user

TABLE 12: "IT GOVERNANCE ARCHETYPES"[238]

In addition, Davenport et al. identify anarchy, Feudalism, Monarchy, and Federalism, as potential governance models.[239]

Governance models are represented graphically in Figure 18 below.

[238] Reproduced from Weill and Ross *IT Governance*, 2004, MIT Sloan School Center for Information Systems Research (CISR), 2003. Cited with permission in Weill and Ross, *IT Governance*, 59.

[239] Davenport et al., "Information Politics," 56.

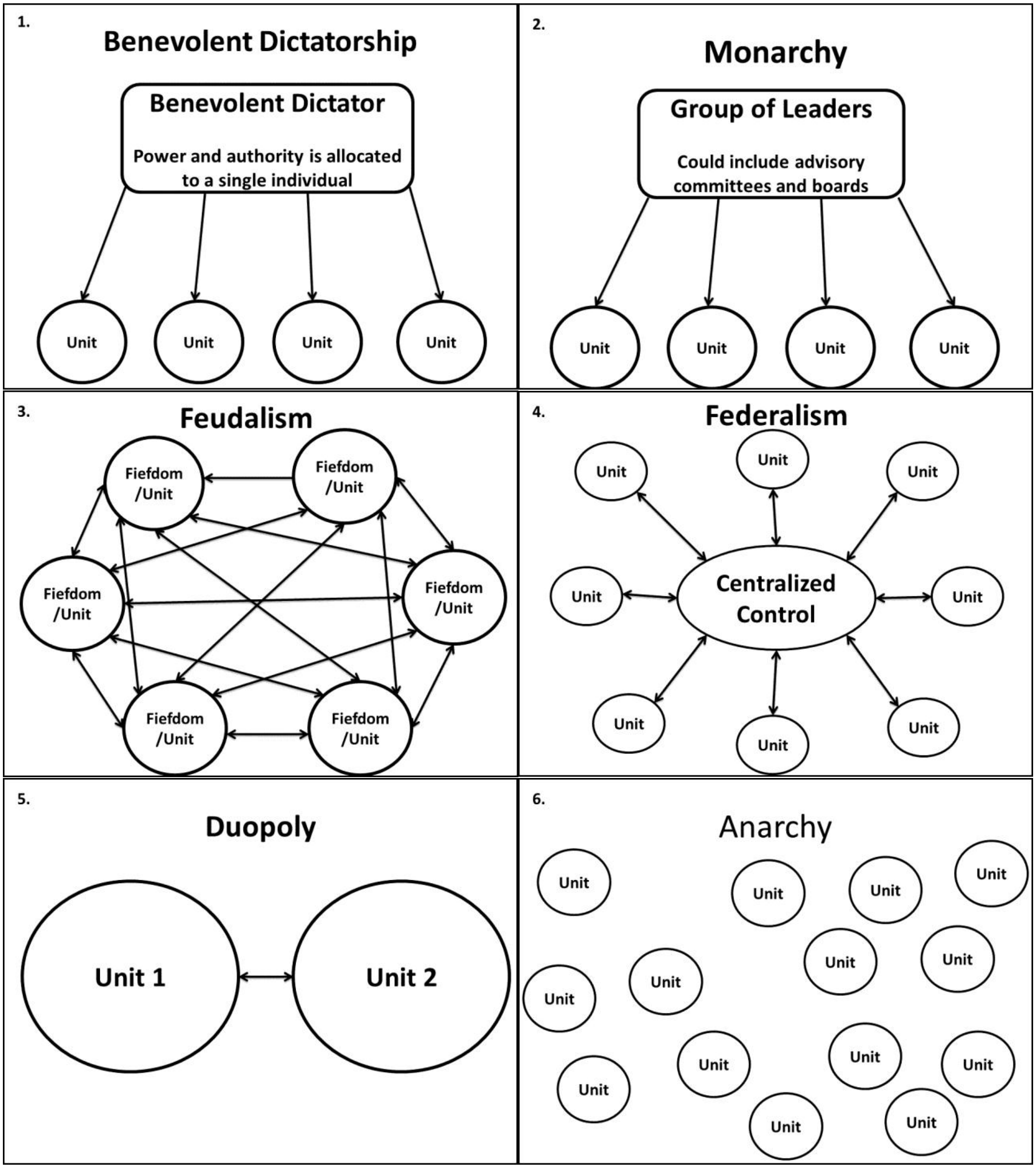

FIGURE 18: GRAPHICAL REPRESENTATIONS OF SEVERAL GOVERNANCE MODELS, ORDERED ON A SCALE OF MOST CENTRALIZED (BENEVOLENT DICTATORSHIP) TO MOST DECENTRALIZED (ANARCHY).[240]

Monarchy. Power is centralized and the organization is structured into information categories and units that report to the firm's leaders. Departments and divisions generally have little autonomy vis-à-vis the firm's information policies, and firm leaders may or may not share collected information with the rest of the firm

[240] Business and IT Monarchy models are represented simply as Monarchy. Meritocracy can be employed to select groups or individuals to make decisions in any one of these models, therefore it is not represented as a stand-alone governance model.

(depending on the "monarch's" approach to information management). A disadvantage of the Monarchy model is that business leader turnover can greatly disrupt information management processes and politics.[241]

Davenport et al. suggest a constitutional monarchy as a solution to limit the monarch's absolute power and achieve greater input from outside the central leadership. A constitutional monarchy is based on a document (constitution) that states the monarchy's limitations, subjects' rights and the law's authority, thereby setting limitations on a monarch's power. This "constitution" states what information will be collected, in what form, by whom, and for what means. The document can also promote the use of a common vocabulary. [242]

Weill and Ross divide the Monarchy model into business and IT. In a Business Monarchy, decisions are made by business executives, individual executives, or committees of senior business executives. Business decisions affect the enterprise as a whole, thus input for decision-making is gathered from many sources. In an IT Monarchy, individuals or groups of IT executives make decisions about IT architecture.[243]

Feudal. Decisions are made by business units, key process owners, or their delegates.[244] Each business unit defines what information will be collected, how it will be interpreted, and in what format it will be reported. Each unit reports limited information to the overall corporation, leading to, "feudal actions [that] diminish the central authority's power to make informed decisions for the common good."[245]

Feudalism is the least effective political model[246] and does not facilitate business-wide decision-making or cooperation.[247] In a Feudal system, individual business units will often restrict access to and distribution of information, thereby reducing the degree of cooperation among individual business units. "The existence of strong, independent, and often warring fiefdoms prevents the development of a common vocabulary and shared meaning."[248] Davenport et al. argue that organizational leaders can foster some degree of cooperation in a Feudal system by creating strategic alliances and common architectures to share information across the organization.

Federal. A Federal system negotiates the interests of central and individual bodies to reach a decision, a process similar to that of state and federal governments working together. The negotiation process brings competing and non-competing parties together, and Davenport et al. found it to be the preferred model for many businesses. A Federalist model requires strong central leadership, and an organizational culture that encourages cooperation and learning. It also requires an astute information manager who understands the value of the information itself and the technology that generated the information and who is able to negotiate among the distributed information centers within the organization. By facilitating contracts between each individual information unit and organization executives to share assets, Davenport et al. argue that a strong information manager can leverage a Federalist system to facilitate a shared understanding of the organization's information vision. [249]

In contrast, Weill and Ross argue that the Federal system is the most difficult decision-making model because business leaders tend to follow their own interests, which may be different from those of other business leaders. A Federal model gives room for bigger, more powerful business units to get the most attention and have the most influence in the decision-making process. It is possible to reduce the effect of unequal power and representation in

---

[241] Ibid.
[242] Ibid.
[243] Weill and Ross, *IT Governance*, 59
[244] Weill and Ross, *IT Governance*, 60
[245] Davenport et al., "Information Politics," 57
[246] Ibid.," 61
[247] Weill and Ross, *IT Governance, 60*
[248] Davenport et al., "Information Politics," 61
[249] Ibid.," 60

the Federal model by establishing management teams and executive committees to act as moderators[250] (similar to the information manager mentioned by Davenport et al.).

IT Duopoly. An IT Duopoly is formed via a bilateral agreement between IT executives and one other group. Weill and Ross argue that this model is advantageous for many reasons. Two decision-making parties can achieve Federal model objectives but with a simpler management structure, and central IT decision-makers who view the enterprise as a whole can identify potential IT and business opportunities. An IT Duopoly can also help manage enterprise-wide adherence to IT architecture. A disadvantage of the IT Duopoly is that it can be expensive and ineffective in dealing with organization-wide issues.[251]

Anarchy. The Anarchy governance archetype is characterized by a lack of an over-arching information management policy, in which individual users make local decisions based on their own needs. This model is very expensive to support and secure.[252] The quality of information collected is likely to be high, but drawbacks of the anarchy model include redundancies, discrepancies, and inefficiencies in information processing and storage.[253]

Governance as a System:

Governance as a system answers the question*: How will these decisions be made and monitored?*

Weill and Ross argue that one governance archetype may not be appropriate for all decisions. Instead, they view governance as a system, composed of many (potentially overlapping) archetypes used, as needed, to generate information and make the types of decisions described above (see page 10). In order to determine the governance archetype best suited to gather information and make each type of decision listed above, Weill and Ross created the governance matrix reproduced below (Table 13). They then conducted a series of case studies on corporate IT governance, mapping the governance archetype to the different types of IT-related decisions. A summary of their results is discussed below in 'How Enterprises Govern.'

Decision **Archetype**	***IT Principles***	***IT Architecture***	***IT Infrastructure Strategies***	***Application Needs***	***IT Investment***
Business Monarchy					
IT Monarchy					
Feudal					
Federal					
Duopoly					
Anarchy					

TABLE 13: "GOVERNANCE AS A SYSTEM ARRANGEMENTS MATRIX—WHICH GOVERNANCE ARCHETYPES ARE USED FOR DIFFERENT TYPES OF DECISIONS?"[254]

[250] Weill and Ross, *IT Governance, 61*
[251] Ibid., 61-63
[252] Davenport et al., "Information Politics," 63
[253] Ibid., 56-57.
[254] Reproduced from Weill and Ross *IT Governance*, 2004, 11. MIT Sloan School Center for Information Systems Research (CISR), 2003. Cited with permission in Weill and Ross, *IT Governance.*

GOVERNANCE AND MANAGEMENT

Weill and Ross argue that governance determines who makes decisions and defines decision-making procedures and that management is the process of making and implementing decisions. They identify four key management principles:[255]

1. **Make tough choices.** Strategic priorities must be established to give clear direction, and there must be a commitment to a small set of the most critical objectives. Tradeoffs are necessary and lead to tough choices.
2. **Develop metrics to formalize strategic choices.** Metrics that formalize strategic choices allow IT value to be measured, making it possible to monitor the governance framework in terms of progress toward strategic goals.
3. **Determine where organizational structure limits desirable behaviors and design governance mechanisms to overcome those limitations.** Governance can transcend organizational structures to enable objectives consistent with organization design.
4. **Allow governance to evolve as management learns role of IT and how to accept accountability for maximizing IT value.**

GOVERNANCE MECHANISMS

The governance models discussed above are implemented on a day-to-day basis through governance mechanisms, including decision-making structures, alignment processes, and communication processes.[256] Weill and Ross suggest effective governance mechanisms must be aligned with strategy, structure, and desired outcome, arguing that "IT governance can be messy. Governance fosters debate, negotiation, constructive disagreement, mutual education, and often frustration. The process is messy, but good governance arrangements enable individuals representing an enterprise's conflicting goals to reconcile their views to the enterprise's benefit."[257]

Decision-Making Structures: Organizational units and roles responsible for making IT decisions, including committees, executive teams, and business/IT relationship managers.

Alignment Processes: Formal processes that ensure behaviors are consistent with IT policies and provide input for IT decisions. They include IT investment proposal and evaluation processes, architecture exception processes, service-level agreements, chargeback, and metrics.

Communication Processes: Announcements, advocacy, channels, and education efforts that inform the organization of IT governance principles and policies and outcomes of IT decision-making processes.

Effective governance mechanisms should be simple, transparent, and suitable, and should "unambiguously define the responsibility or objective for a specific person or group."[258] Individuals should be engaged in the best way to make given decisions, and individuals affected by, or who want to challenge, the governance process should be aware of how the mechanisms work.

Weill and Ross recommend five principles for designing effective governance mechanisms:[259]

1. **Choose mechanisms from all three types:** Decision-making, alignment, and communication mechanisms have different objectives, but all are important for effective governance.

[255] Weill and Ross, *IT Governance*, 183-184.
[256] Ibid., 86
[257] Ibid., 85.
[258] Ibid., 114-155.
[259] Ibid., 115-116

2. **Limit decision-making structures:** Complex organizations need multiple decision-making structures, but too many decision-making structures give room to contradictions and communication gaps. Alignment mechanisms should disseminate decision-making responsibilities throughout the organization.
3. **Provide for overlapping membership in decision-making structures:** Key decision-making groups need overlapping membership or clear mandates to avoid disconnects between IT and business direction.
4. **Implement mechanisms at multiple levels in the organization:** Different levels have different needs and objectives, so governance should match accordingly. Architecture committees and IT budget processes can serve as essential connections between the firm as a whole and individual business units.
5. **Clarify accountability:** Clear management objectives and metrics can help avoid confusion over who is responsible for what.

EXAMPLES OF GOVERNANCE

It is instructive at the early development stages of EarthCube to consider how governance is employed in a variety of different fields and organizational structures. This section reviews IT governance, non-profit governance, and open source software governance.

CORPORATE IT GOVERNANCE

There are four components of corporate IT governance: strategy, desirable behaviors, assets, and governance of assets[260]. Strategy refers to a set of choices. In corporations, these choices refer to targeted customers, products and service offerings, etc. Desirable behaviors encompass the beliefs and culture of the organization defined and enacted by strategy and mission statements, etc. Assets include the following:[261]

- **Human:** people, skills, training, mentoring, competencies, reporting, career paths
- **Financial:** cash, investments, liabilities, receivables, cash flow
- **Physical:** buildings, plants, equipment, maintenance, security
- **Intellectual Property:** product services, patents
- **Information and IT:** digitized data, information systems
- **Relationships:** within the enterprise and without

HOW ENTERPRISES GOVERN

Weill and Ross found a number of common governance patterns in their research. Firms gathered input from many sources and decision-making rights were allocated to different groups, depending on the decision being made. The Federal model was most common for business-oriented decisions, and committees, budgets, and cross-functional process teams to provide input and feedback were common governance mechanisms. The IT Duopoly was often used for less technical decision-making, focusing on decisions regarding IT principles, business application needs, and IT investment, for example, but technical decisions were commonly made via an IT Duopoly and Federal System.

Weill and Ross' findings are summarized in Table 14 below. Each type of decision is divided into two processes: 1) the process employed for gathering information regarding the decision at hand and 2) the process for actually making the decision.

[260] Weill and Ross, *IT Governance*, 6
[261] Ibid, 6-7

How Enterprises Govern

DECISION

GOVERNANCE ARCHETYPE	*IT Principles*		*IT Architecture*		*IT Infrastructure Strategies*		*Business Application Needs*		*IT Investment*	
	Input	Decision	Input	Decision	Input	Decision	Input	Decision	Input	Decision
Business Monarchy	0	27	0	6	0	7	1	12	1	30
IT Monarchy	1	18	20	73	10	59	0	8	0	9
Feudal	0	3	0	0	1	2	1	18	0	3
Federal	83	14	46	4	59	6	81	30	93	27
Duopoly	15	36	34	15	30	23	17	27	6	30
Anarchy	0	0	0	1	0	1	0	3	0	1
No Data or Don't Know	1	2	0	1	0	2	0	2	0	0

Most common input pattern for all enterprises.

Most common decision patterns for all enterprises.

The numbers in each cell are percentages of the 256 enterprises studied in twenty-three countries. The columns add to 100 percent.

TABLE 14: HOW ENTERPRISES GOVERN[262]

CHOOSING THE RIGHT GOVERNANCE MODEL

Weill and Ross recommend the Federal model is best for gathering input for IT decisions, and the Duopoly model is best for IT principles and investments-related decisions. They suggest avoiding the Federal model for decision-making, but if the Federal model is used, there should be mechanisms in place to balance the needs of the center and surrounding business units. For any decision regarding investments, principles, and business application needs, joint-decision making involving business and IT representatives is best.[263]

In contrast, Davenport et al. argue that Monarchy and Federalism are the only viable choices for an information governance model. Federalism is preferable in an organizational culture of broad participation and empowerment, although a Federalist model can be difficult to implement. The benevolent Monarchy model is aptly suited for firms that experience difficulties in achieving consensus. A strong, top-down approach ensures a common vocabulary and supports autonomy and coordination, while the right processes create information aligned with firm objectives. Davenport et al. offer a number of additional suggestions for choosing the right governance model to achieve desired outcomes of the organization or firm. Specifically, they recommend matching the organizational politics to the organizational culture, practicing technological realism, and electing the right information politicians.[264]

---

[262] Reproduced from Weill and Ross, *IT Governance*, 64. MIT Sloan School Center for Information Systems Research (CISR), 2003. Cited with permission in Weill and Ross, *IT Governance.*
[263] Weill and Ross, *IT Governance*, 129-133.
[264] Davenport et al., "Information Politics," 60-64

Select an Information State. The first step in choosing a governance model is to examine what information management models people in the firm already hold, which model is most dominant, which is most desirable, and how it can be achieved. Davenport et al. recommend choosing one model and moving toward it, no matter how long it takes, because multiple models will create confusions and waste resources. Assessment tools for evaluating a particular model's effectiveness include evaluating commonality of vocabulary and meaning, degree of access to important information, efficiency of information management, and quality of information in terms of currency, relevance and accuracy,

Match Information Politics to Organizational Culture. More democratic information-management practices will only occur when a firm's culture is accepting of the processes that lead to democracy. Davenport et al. argue that "a firm's culture must be conducive to participative information management and free information flow before they will happen. Put another way, information flow does not make an organizational culture less hierarchical and more open; rather, democratic culture makes possible democratic information flows."[265] Indicators that a firm is ready for more democratic information management practices include empowerment of workers to make decisions, collaborative work to improve organization processes, and removal to the greatest extent possible of fear as a motivator.

Practice Technological Realism. IT engineering should be focused on concrete goals or processes, including technological elements that are crucial to strategy implementation and day-to-day functions. Information should also be presented in units that managers can understand and negotiate with, such as forms, reports and memo, instead of whole datasets. Technology platforms should be easy use-to-use, and be based on common, standardized technology.

Elect the Right Information Politicians. The primary role of individuals or groups in charge of information management should be to facilitate the use of information across the organization. These people should be able to convince others of the importance of information management, and the reasons why the chosen model is the best model for that firm. The responsibility for collecting, maintaining, and interpreting information should not be placed in the hands of just one person. Instead, specific information management roles should be assigned throughout the organization.

CHARACTERISTICS OF TOP CORPORATE GOVERNANCE PERFORMERS

Weill and Ross found many common qualities among governance top-performers. Top-performing firms had governance structures linked to performance measure on which they excelled (for example, growth or return on assets). Top performers also aligned business objectives with governance approaches, governance mechanisms, and performance goals and metrics.[266] Finally, top performers balanced multiple performance goals, their governance models blended centralized and decentralized decision-making, and the IT decision-making process, such as resolution of the tension between standardization and innovation, was made transparent.[267]
Weill and Ross also attribute governance performance with the specific archetypes used for gathering input and making decisions. Top governance performers used the Federal system to gather input about IT principles and business application needs, but used the Duopoly archetype to actually make decision (Table 15).

[265] Ibid.,62.
[266] Ibid., 115
[267] Weill and Ross, *IT Governance*, 18.

Best and Worst Governance Performers Use Different Arrangements

DECISION

GOVERNANCE ARCHETYPE	*IT Principles*		*IT Architecture*		*IT Infrastructure Strategies*		*Business Application Needs*		*IT Investment*	
	Input	Decision	Input	Decision	Input	Decision	Input	Decision	Input	Decision
Business Monarchy										
IT Monarchy										
Feudal								-		
Federal	+	-		-		-	+			
Duopoly	-	+					-			+
Anarchy										

- Poor performers   + Top performers

TABLE 15: BEST AND WORST GOVERNANCE PERFORMERS USE DIFFERENT ARRANGEMENTS[268]

Top governance performers had seven characteristics in common: [269]

1. More managers in leadership positions are able to describe the organization's IT governance
2. Senior management is engaged in governance and communicates to the rest of the organization
3. More direct involvement of senior leaders in IT governance
4. Clear business objectives for IT investment
5. Differentiated business strategies
6. Fewer renegade decisions and more formally approved exceptions
7. Fewer changes in governance from year to year

Weill and Ross highlight the top 10 leadership principles of effective governance: [270]

1. Actively design governance
2. Know when to redesign
3. Involve senior managers
4. Make choices
5. Clarify the exception-handling process
6. Provide the right incentives
7. Assign ownership and accountability for IT governance
8. Design governance at multiple organizational levels
9. Provide transparency and education

[268] MIT Sloan School Center for Information Systems Research (CISR), 2003. reproduced from Weill and Ross, *IT Governance*, 131.
[269] Weill and Ross, *IT Governance*, 145-146
[270] Ibid., 222-230

10. Implement common mechanisms across the six key assets

Weill and Ross also note six interlocking components of effective governance design: [271]

1. Enterprise strategy and organization
2. IT organization and desirable behavior
3. IT governance arrangements (combination of governance archetypes used to make IT decisions)
4. IT governance mechanisms
5. Business and performance goals
6. IT metrics and accountabilities

These elements must be harmonized to meet organizational goals. Specifically, Weill and Ross argue that:

1. Enterprise strategy and organization must be harmonized with IT organization and desirable behavior
2. IT governance arrangements (decision rights via monarchies, Federal, etc.) must be harmonized with IT governance mechanisms (committees, budgets, etc.) to effectively make IT decisions on principles, architecture, infrastructure, applications and investment
3. Business and performance goals must be harmonized with IT metrics and accountabilities.

Additionally, all six components must be connected so that the overall governance framework advances organizational goals.

NON-PROFIT GOVERNANCE

Weill and Ross found many similarities between non-profit and corporate IT governance, but with a few key distinctions. They note that different approaches come from a broader definition of public value and a culture that values consensus, transparency, and equity. Challenges faced by not-for-profits include difficulties in measuring performance and value, a culture of formal committees, limited budgets, endless opportunities to create value, and the fact that performance measures are not as clear as in for-profit firms.[272]

Weill and Ross recommend forming governance structures around a value framework based on authorizing environment (potential clients or customers, funding sources and political influences), capabilities (resources in the form of money and authority to create capabilities), and public value (generated in addition to private value, includes public goods and equity, and extends beyond the organization). Public value created by the not-for-profit goes back to the authorizing environment in the form of transfers and provision of services, and the authorizing environment grants permission back to the creators of public values in the form of market signals (increased demand) and political acceptance.

Weill and Ross argue that the authorizing framework, capabilities, and public value must be aligned toward the goal of creating public value to the maximum extent possible given the authorizing environment expectations and available internal and external capabilities.

TOP NOT-FOR-PROFIT GOVERNANCE PERFORMERS

Weill and Ross encountered key characteristics of successful governance in not-for-profits. These include partnerships and joint decisions between IT and not-for-profit leaders, and heavy use of mechanisms, such as formal committees, and a broad definition of value (recognized by the inclusion of representatives from outside the not-for-profit in IT governance mechanisms). Stable governance mechanisms that change with less frequency are

[271] Weill and Ross, *IT Governance*, 148 - 157
[272] Ibid., 195 - 200

also a key characteristic because communication and implementation of new procedures take longer than in for-profit firms. Weill and Ross offer a number of suggestions for successful not-for-profit governance:[273]

1. Executive committees focused on all key assets including IT
2. IT council comprising business and IT executives
3. IT leadership committee comprising IT executives
4. Architecture committee
5. Tracking of IT projects and resources consumed
6. Business/IT relationship managers

OPEN SOURCE GOVERNANCE

Since EarthCube is envisioned to contain software development, Open Source Governance is also included in this review. In an open source software project, the "governance model describes the roles that project participants can take on and the process for decision making within the project. In addition, it describes the ground rules for participation in the project and the processes for communicating and sharing within the project team and community."[274]

Effective development and communication of a clear governance model is an important aspect of any sustainable project. A governance framework also creates quality control processes that assure potential users of project viability. It also communicates policies and strategies to potential users and developers of project's outputs. In this way, potential contributors understand how they can contribute, what is expected of them, and how their contributions will be handled.

OPEN SOURCE SOFTWARE GOVERNANCE MODELS

OSS Watch suggests a benevolent dictatorship, a meritocracy, or some combination of the two as possible models for open source software projects, although it is possible for the governance model to evolve as the project matures. In a benevolent dictatorship, a single individual or organization maintains centralized control. Project founders retain control throughout the life of the project, determining strategic direction and maintaining the authority to make decisions, particularly when the community is in dispute. Community discussion and debate should inform the decision-making process, and decisions should be made in the best interests of the organization.

In a meritocracy, distributed control is awarded in recognition of contributions to the project, and reward mechanisms are built into the governance framework. Project leaders have earned the respect and recognition of the community, and everyone can contribute to the decision-making process. Structured decision-making processes are required because no single individual can break a community-deadlock.

EXAMPLES OF OPEN SOURCE SOFTWARE GOVERNANCE

Ubunto, Apache Software Foundation, and Taverna Project are examples of existing open source governance models.[275] The Ubunto governance model outlines decision-making processes, provides clear documentation of technical management processes, and describes the community structure and responsibilities that go with each component of that structure. The Apache Software Foundation employs two sets of governance documents for each project. The first document lays out the governance framework of the Apache Software Foundation as a whole, including information on key processes within each project, such as decision-making. The second document is

[273] Weill and Ross, *IT Governance, 205-206*
[274] Ross Gardler and Gabriel Hanganu, "Governance Models"
[275] Ibid.

created by each individual project, within certain parameters. This process allows each project to define its own variations of Foundation-wide governance processes. The Taverna Workflow Management System is an example of an open source project used by more than 350 academic and commercial organizations worldwide. Taverna Governance framework focuses on project structure and management processes and is an example of a model that can work in an academic environment. Existing open source governance models are displayed below (Table 16).

Open Source Software Models

Project type	Objective	Control style	Community structure	Examples
Exploration-oriented	Sharing innovation and knowledge	Cathedral-like central control	Project leader, many readers	Perl, Linux Kernel
Utility-oriented	Satisfying an individual need	Bazaar-like decentralized control	Many peripheral developers, peer support to passive users	GNU/Linux (excluding Kernel)
Service-oriented	Providing stable service	Council-like central control	Core members instead of project leader, many users that develop systems for end users	Apache, PostgreSQL

TABLE 16: COMMON OPEN SOFTWARE GOVERNANCE MODELS[276]

SUSTAINABILITY IN OPEN SOURCE SOFTWARE PROJECTS

Although Hanganu and Gardler write about sustainability in relation to open source software governance, sustainability is a key factor in the success of any CI project and is an important component of any governance framework. A sustainability plan must be in place from the beginning of the project and must be maintained throughout the life of the project. Hanganu and Gardler list the key sustainability issues in open source software projects as: 1) Failure to attract interest from third-parties; 2) Poor Management of community infrastructure; 3) Lack of Project Memory; and 4) Inefficient Release Management.[277] Although these issues are examined in relation to open source software projects, they may be relevant to EarthCube.

Failure to attract third-party interest. Two key questions must be asked when determining the viability of an open source software project: "Who is the project community? Do I need to build a new community, or is there an existing one my project should be part of?"[278] Thus, the first step of any open source software project is to investigate other projects and examine whether these projects can be leveraged, or if the proposed project can contribute to existing work. This process avoids duplication, and links the project to the larger community, thus increasing project sustainability.

Poor management of community infrastructure. The project should clearly indicate the type of license that governs the software output. There should be a "clear and unambiguous record of all project contributions,"[279] which can be simplified using a version control system as part of open development tools and supporting processes. Documentation is also a key component of good community infrastructure management: "A true open source project ensures that all the solutions adopted, including the decision-making process used in reaching them, are openly shared and appropriately recorded for later reference. This is because the reasons behind a particular decision are often more important, with respect to project sustainability, than the software itself. Such

[276] Ross Gardler and Gabriel Hanganu, "Governance Models."
[277] Ibid.
[278] Ibid..
[279] Ibid.

documentation reduces the chances of repeating past mistakes, and ensures that old ground does not need to be covered repeatedly."[280]

Lack of project memory. Projects need tools to preserve project memory that allow all contributors to understand the project context and motivations for past decisions:

> *"Project memory needs to be created for those who come after the initial funded effort. Without community infrastructure there is no project memory; without memory the only way to keep the project alive is to get more funding for the originating team. Failing to set up project memory during funded development will affect the creation of new projects based on the old. It also makes it much harder for external potential collaborators to evaluate your project and fully understand the project."*[281]

Project memory should be openly maintained via a mailing list that documents discussions and decisions, an issue tracker to record and prioritize issues, and end-user feedback. This process creates project memory for new users, which is a key component in ensuring project sustainability.

Inefficient release management. Software should be released early and often, and there should be a well-defined process to manage project releases. A well-defined release management process helps developers coordinate to produce new versions of software. Hanganu and Gardler argue that the technical details of release should be separate from the managerial release components. These and other sustainability issues and governance mechanisms are summarized below (Table 17):

Sustainability Resources

Title	Hours	Activity	Benefits
Governance model	30	Understand the different governance models. Take a governance document template from OSS Watch and customise it for the project.	Defines scope of project and decision-making process for strategic management
Infrastructure set-up	10-30	Set up, configure and document the tools needed for the project. Quicker if using a public hosting provider like Sourceforge or Googlecode, longer if using local hosting facilities.	Streamlines the software management aspects of the project and channels community members into the management process
Project memory	8 per month per full time contributor	Document all activity in a visible way. Ensure all project communication is recorded on mailing lists and all proposals are brought to the lists for discussion and approval. Record and prioritize issues in the issue tracker.	Records project memory and encourages third-party engagement
Release management	8 per month	Set up a release management process and ensure all developers follow it. Develop and maintain release build scripts. Ensure all legal issues are addressed and the quality of each release is good enough for users to install and run easily.	Manages the "release early, release often" requirement for sustainable open source. Facilitates third-party engagement by ensuring they can build the project software easily

TABLE 17: COMMON SUSTAINABILITY RESOURCES IN OPEN SOURCE SOFTWARE PROJECTS.[282]

[280] Ibid.
[281] Ibid.
[282] Ross Gardler and Gabriel Hanganu, "Governance Models."

ROLES AND RESPONSIBILTIES IN OPEN SOURCE SOFTWARE PROJECTS

An important component of any governance framework is defining the roles within the organization or project and determining processes by which those roles may be filled. Open source software projects provide many examples of potential project roles. Elizabeth Tatham of OSS Watch argues that:

> *"You don't need to be a software developer to contribute to an open source project. The code, documentation and artwork that make up an open source project have all been created, tested, used, discussed and refined by members of the project community. These processes can be broken down into a myriad tasks, requiring different skills, levels of involvement and degrees of technical expertise. So, if you want to get involved in an open source project, there is a range of roles to choose from."*[283]

Possible tasks include providing feedback, helping new users, recommending the project to others, testing and reporting or fixing bugs, requesting new features, writing and updating software, creating artwork, writing or updating documentation, and translating.[284]

BARRIERS TO ADOPTING A GOVERNANCE MODEL IN OPEN SOURCE SOFTWARE PROJECTS

OSS Watch cites a number of common barriers to adopting a governance model. These barriers include the perception that the governance process is just 'red tape,' concerns that a governance framework may cause the project to lose its sense of direction, fears that control of the project will be lost, and a belief that the project is too young or too small to attract active users or developers. The first three concerns are valid but can be remediated with a proper governance model. The fourth concern (belief that the project is too young or too small to warrant a governance model) is never valid, however, because potential contributors will always need clear guidance about how their contributions will be handled, and projects cannot afford to lose contributors at the beginning. These concerns must be addressed at the beginning of the project to avoid becoming a barrier to adopting a governance model later on.

HOW TO CREATE A GOVERNANCE DOCUMENT FOR OPEN SOURCE SOFTWARE PROJECTS

OSS Watch offers a number of recommendations for how to draw up a governance model for an open source software project. The document should outline decision-making processes and explain how contributions will be handled. The initial governance framework does not have to be very detailed, and every potential situation that may arise in the future does not have to be accounted for. Instead, the model should be easily understandable and manageable, and should send a clear message to potential project contributors. The document should be concise, accessible, and easy to reference to ensure that it is used throughout the lifespan of the project.

Overview. An overview of the governance document should include a brief summary of project objectives, how the project will be managed, how potential contributors can become involved, identification of the copyright holder(s), and explanation of the licenses covering project outputs.

Roles and Responsibilities. The formal roles and authority within the project should be established in the governance document, specifying the level of commitment required and how individuals can take on different roles. These roles can be general or specific and should strike a balance between too many specific roles (which would limit contributions) and not enough direction (which can lead to a lack of contribution).

---

283 Elizabeth Tatham, "Roles in Open Source Projects," Open Source Software Advisory Service, last modified May 15, 2012, http://www.oss-watch.ac.uk/resources/rolesinopensource.xml
284 Elizabeth Tatham, "Roles in Open Source Projects."

Support. Support processes within the project (engagement of end-users, for example) should be defined, including available means of support and how each one may be used. This will prevent spreading support too thin, which may limit the ability of the community to be self-supportive.

Decision-Making Processes. Openness and transparency in the decision-making processes should be emphasized. There should be a focus on striking a balance between decision-making processes being too time-consuming and not being well-defined.

IMPORTANCE OF GOVERNANCE

Governance can help assure that decisions throughout the enterprise are consistent with the organizational direction and can, in fact, influence how that direction is established. Effective governance aligns daily behaviors, decision-making processes, and communication mechanisms to meet overall organizational goals.

CORPORATION IT GOVERNANCE

Good governance is important in corporations for many reasons:[285]

Good IT governance pays off. Good IT governance leads to superior profits because, "well-defined IT governance arrangements distribute IT decision-making to those responsible for outcomes."[286]

IT is Expensive. Focused IT spending on strategic priorities can reduce IT costs by involving the right people in the IT decision-making process.

IT is pervasive. "Well designed IT governance arrangements distribute IT decision making to those responsible for outcomes."[287]

New Information Technologies Bombard Enterprises with New Business Opportunities. Governance can also take advantage of opportunities created by new technologies by having the foresight to establish the right infrastructure at the right time to create new business opportunities.

IT Governance Is Critical to Organizational Learning About IT Value. Effective governance creates mechanisms to debate the potential value of specific IT initiatives and formalize learning about IT and thus is critical for the organization to learn about IT value.

IT Value depends on more than good technology. "Successful firms involve the right people in the process. Having the right people involved in IT decision making yields both more strategic applications and greater buy-in. These more involved people then produce better implementations."[288]

Senior management has limited bandwidth. "Decisions throughout the enterprise should be consistent with the direction in which senior management is taking the organization. Carefully designed IT governance provides a clear, transparent IT decision-making process that leads to consistent behavior linked back to the senior management vision while empowering everyone's creativity."

Leading enterprises govern IT differently. "Top performing firms balancing multiple performance goals had governance models that blended centralized and decentralized decision making. All top performers' governance had one aspect in common. Their governance made transparent the tensions around IT decisions such as standardization versus innovation."[289]

INFORMATION MANAGEMENT

Governance also improves information management.[290] Information politics play an increasingly important role in modern for-profit and not-for-profit organizations, universities, virtual organizations, and cyberinfrastructure (CI) projects. As jobs become defined by the information they possess, people may be less likely to share that

---

285 Weill and Ross 2004, *IT Governance:* 14-18.
286 Ibid., 15
287 Ibid., 15
288 Ibid., 17
289 Ibid., 18.
290 Thomas Davenport, Robert Eccles, and Laurence Prusak, "Information Politics," 1992.

information. A governance framework allows firms, organizations, universities and other entities to address not only the information itself, but the organizational structure that leads to the generation of that information, and the processes that determine how that information is handled.[291]

OPEN SOURCE GOVERNANCE

Effective governance in open source software projects also has a series of benefits.[292]

A governance framework should be simple, but effective. It is not necessary to create rules for every situation, particularly at the beginning of the project. Open Source Software (OSS) Watch states, "the model should provide a lightweight way of guiding the management of the project in a clear and transparent way. This transparency will encourage third parties to join the project. They can see how the project is run and predict how it will react to their contributions before expending any significant effort on that work."[293]

An effective governance framework increases the agility of the project. A good governance framework can be flexible while maintaining core goals. This can be achieved through governance mechanisms that allow for community input while project leaders maintain overall control, effectively using time and resources to make the process of evaluating potential contributions easier.

An effective governance framework increases project sustainability. The goal of a governance framework should be to provide complete solutions to all stakeholders while making sure that the project is still of interest to potential contributors. The governance framework should build in mechanisms to accommodate future needs and potential changes in project direction. In this way, the governance model increases the resource pool that can be used to maintain project sustainability.

An effective governance framework helps maintain a clear sense of direction. A governance framework should have mechanisms that define and enforce boundaries, which leads to time saved by avoiding or minimizing rogue elements. Groups and individuals with a clear sense of direction and aligned elements can more easily collaborate and complement each other's work, which broadens the scope of the project.

An effective governance framework allows control to remain where the leaders want it to be. Decision-making can be controlled by a central organization or can be the responsibility of the community. The level of control and accompanying decision-making mechanisms depend on the type of project and how efforts are funded.

An effective governance framework encourages potential contributions and project growth. A good governance model clearly demonstrates how potential contributions will be managed, thus encouraging greater contributions from those outside of the original project community. Every contribution can improve the quality of the overall project, and every potential contributor who decides not to contribute is a missed opportunity for project's growth and sustainability.

Contribution Process. How people should contribute to the process should be specified, including the types of contributions solicited, and the tools available to manage and make contributions. This guidance acts as a mechanism for quality control and states expectations regarding copyrights, coding standards and documentation. It also describes the process for third party contributions and what happens when a contribution is deemed unacceptable.

[291] Ibid.
[292] Ross Gardler and Gabriel Hanganu, "Governance Models."
[293] Ibid.

Governance Definitions:

- IT governance "specif[ies] the decision rights and accountability framework to encourage desirable behavior in using IT. IT governance is not about making specific IT decisions–management does that–but rather determines who systematically makes and contributes to those decisions."[294]
- Governance "provid[es] the structure for determining organizational objectives and monitoring performance to ensure that objectives are attained."[295]
- A "governance model describes the roles that project participants can take on and the process for decision making within the project. In addition, it describes the ground rules for participation in the project and the processes for communicating and sharing within the project team and community."[296]
- Governance "aligns an organization's practices and procedures with its goals, purposes, and values. Definitions vary, but in general governance involves overseeing, steering, and articulating organizational norms and processes (as opposed to managerial activities such as detailed planning and allocation of effort). Styles of governance range from authoritarian to communalist to anarchical, each with advantages and drawbacks."[297]

Types of Governance Reviewed:

Corporate IT governance
Non-profit governance
Open Source Software Governance

Effective governance must address three key questions:

What decisions must be made to ensure effective management and use of IT?
Who should make these decisions?
How will these decisions be made and monitored?

Five key types of decisions that must be made:

IT principles: Clarifying the business role of IT
IT architecture: Defining integration and standardization
IT infrastructure: Determining shared and enabling services
Business application needs: Specifying business need for purchased or internally developed IT applications
IT Investment and prioritization: Choosing which initiatives to fund and how much to spend

Governance archetypes that determine who makes decisions and how decisions are made:

Business Monarchy: top managers
IT Monarchy: IT specialists
Feudal: each unit makes independent decisions
Federal: combination of corporate center and business units with or without IT people involved
IT Duopoly: IT group and one other group (ex: top management or business unit leaders)
Anarchy: isolated individual or small group decision making
Meritocracy: distributed control is awarded in recognition of contributions to the project
Technical Utopianism: technical approach to information management.

---

294 Weill and Ross, *IT Governance:* 2.
295 Organization for Economic Cooperation and Development, 5, 219, quoted in Weill and Ross, *IT Governance*, 4-5
296 Ross Gardler and Gabriel Hanganu, "Governance Models."
297 "Governance," EarthSystem Commodity Governance Project, last modified 2012, http://earthsystemcog.org/projects/cog/governance_object

Choosing the Right Governance Model:

Per Weill and Ross

1. Federal model for input into IT decisions
2. Duopoly for IT principles and investment decisions
3. Avoid Federal model for decision-making if possible; If used, model must be mature enough to balance needs of center and business units
4. Duopoly involving business and IT representatives for business-oriented IT decisions (investments, principles, business application needs

Per Davenport et al.

1. Select an information state
2. Match information politics to organizational culture
3. Practice technological realism
4. Elect the right information politicians

Governance mechanisms: implement the governance framework on a day-to-day basis.

1. Decision-making structures
2. Alignment processes
3. Communication processes

Five principles for designing effective governance mechanisms include:

1. Choose mechanisms from all three types (listed above)
2. Limit decision-making structures
3. Provide for overlapping membership in decision-making structures
4. Implement mechanisms at multiple levels in the organization
5. Clarify accountability

Characteristics of Top Corporate Governance Performers

1. More managers in leadership positions are able to describe IT governance
2. Senior management is engaged in governance and communicates to the rest of the organization
3. More direct involvement of senior leaders in IT governance
4. Clear business objectives for IT investment
5. Differentiated business strategies
6. Fewer renegade decisions and more formally approved exceptions
7. Fewer changes in governance from year to year

Characteristics of Top Not-for-Profit Governance Performers

1. Executive committees focused on all key assets including IT
2. IT council comprising business and IT executives
3. IT leadership committee comprising IT executives
4. Architecture committee
5. Tracking of IT projects and resources consumed
6. Business/IT relationship managers

Top 10 leadership principles of effective governance:

1. Actively design governance
2. Know when to redesign
3. Involve senior managers

4. Make choices
5. Clarify the exception-handling process
6. Provide the right incentives
7. Assign ownership and accountability for IT governance
8. Design governance at multiple organizational levels
9. Provide transparency and education
10. Implement common mechanisms across the six key assets

A governance document for open source software projects should include:

1. Overview of project objectives, project management, identification of potential contributors and copyright holders, and licenses covering project outputs.
2. Roles and responsibilities.
3. Support processes within the project.
4. Decision-making processes.
5. Contribution process.

SECTION 2: HISTORICAL INFRASTRUCTURE DEVELOPMENT CASE STUDIES AND LESSONS LEARNED

Past infrastructure projects' successes and failures may better inform the eventual governance framework for EarthCube by providing examples of best practices and lessons learned. While EarthCube is a revolutionary new cyberinfrastructure (CI) project, many of the problems faced by past projects are potentially pertinent to EarthCube today. The better our understanding of lessons learned and best practices from past CI projects, the more likely we will be able to overcome or avoid these problems in the future.

In an effort to more clearly understand how decision making models and governance frameworks can impact the outcome of programs and projects, this section identifies past infrastructure project successes and failures by presenting examples of best practices and lessons learned. While EarthCube is a new community-based cyberinfrastructure initiative for the geosciences, many of the issues faced by past projects can provide valuable insight into successful governance mechanisms. We begin by reviewing "Understanding Infrastructure: Dynamics, Tensions, and Design - Report of a Workshop on 'History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures'"[298] which focuses on common dynamics, tensions, and design issues across historical infrastructure projects and offers recommendations for future projects. The review continues with an analysis of "Mohole: the Project That Went Awry"[299] which examines the complex factors, primarily associated with governance, that led to the failure of Project Mohole in 1966. The historical review ends with a summary of "Collaboration, Alignment and Leadership"[300] a comparative study of leadership principles among National Science Foundation Network (NSFNET), Unidata, Digital Library for Earth System Science (DLESE), and the National Digital Science Library (NSDL).

UNDERSTANDING INFRASTRUCTURE

"Understanding Infrastructure: Dynamics, Tensions, and Design - Report of a Workshop on 'History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures'" compiled by Edwards, Jackson, Bowker and Knobel identifies a series of common trends and patterns across such varied historical infrastructure projects as the US railway system, telephone system, and the creation of the internet. Cyberinfrastructures are composed of global, social, technical and local elements and are defined as "the set of organizational practices, technical and infrastructure, and social norms that collectively provide for the smooth operation of scientific work at a distance. All three are objects of design and engineering; a cyberinfrastructure will fail if anyone is ignored."[301] Edwards et al. argue that CI development is a series of choices between a social and/or technical solution for any given problem, or some combination of the two (Figure 19).

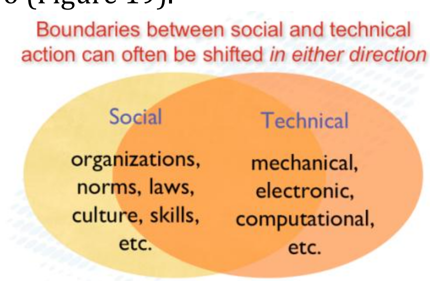

FIGURE19: SOCIAL AND TECHNICAL BOUNDARIES IN CYBERINFRASTRUCTURE[302]

[298] Paul Edwards, Steven Jackson, Geoffrey Bowker, and Cory Knobel, "Understanding Infrastructure: Dynamics, Tensions, and Design - Report of a Workshop on "History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures," 2007.
[299] See D.S. Greenberg, " Mohole: the Project That Went Awry," *Science*, 143, no. 3602, (1964): 115-119,
[300] David Fulker, "Collaboration, Alignment and Leadership, invited piece in 'NDSL Reflections."
[301] Edwards et al., "Understanding Infrastructure," 6.
[302] Ibid., 4.

Edwards et al. found that infrastructure building is usually done by many builders, planning is not always intentional, progress is usually incremental and modular, and the final version of the infrastructure is usually very different from the initial version. They also found that science, theory, and inquiry are created locally, and then brought in as necessary as the infrastructure grows and new communities begin to contribute to the development process.

COMMON STEPS

Edwards et al. identify three common steps in infrastructure development: system building, technology transfer, and integration.[303]

System Building

This phase is carried out by individual, groups, and organizations that create and promote systems, defined as "linked sets of devices that fill a functional need."[304] System building is characterized by a "deliberate and successful design of technology-based services"[305] and occurs when local systems are linked into networks, leading to distributed control and access. When networks are formed, the infrastructure creation process is initiated; "true infrastructures only begin to form when locally constructed, centrally controlled systems are linked into networks and internetworks governed by distributed control and coordination processes."[306]

Technology Transfer

The second phase occurs when technology is transferred from one location or domain to another and the system adapts to new conditions. Competing systems emerge at this time, in addition to the growth in scale of the infrastructure project as a whole. The technology transfer process creates new legal, technical, social, cultural, organization, and financial adjustments, which produce new organizational support and require technical innovation. In addition, the growth in scale of the infrastructure project incorporates many more heterogeneous elements than in the first phase, and finance capital, legal representation, and political and regulatory relationship management become essential components of the system as a whole.

Edwards et al. argue consideration of users and user communities is critical to the success of a CI project, particularly during this phase. They note that, "a key problem is that the development process builds expertise among the developers; as a result, developers can lose their ability to see how novices, or users in a different field, perceive and use their system."[307] As IT projects grow and transition to a more diverse community of novice users, these users often take the process into their own hands, which often results in divergent norms, practices, and standards implementation.[308]

Integration and Consolidation

Heterogeneous systems are linked into networks through gateways, consolidating the infrastructure project. Competition among technological systems and standards is resolved. This step is complete when the service provided by the infrastructure project becomes a commodity resource.

INFRASTRUCTURE DYNAMICS

[303] Edwards et al., "Understanding Infrastructure," 7-14.
[304] Ibid., 8
[305] Ibid., i
[306] Ibid., 7
[307] Ibid., 8
[308] Claude Fischer, *America Calling: A Social History of the Telephone to 1940*, (Berkeley, CA: University of California Press), 1992; Ole Hanseth and Eric Monteiro, *Understanding Information Infrastructure* (unfinished book manuscript), www.ifi.uio.no/~oleha/Publications/bok.html, 1998; Brian Kahin and Janet Abbate, S*tandards Policy for Information Infrastructure*,(Cambridge, MA: MIT Press),1995; and Janet Abbate, *Inventing the Internet,* (Cambridge, MA: MIT Press), 1999, quoted in Edwards et al., "Understanding Infrastructure*,"* 10.

Edwards et al. identify four additional dynamics that shape the infrastructure development outlined above: reverse salients, gateways, path dependence, and scale effects.[309]

Reverse Salients. Reverse salients are critical, unsolved problems, defined as "points where an advancing front is held back."[310] Reverse salients can be technical, social, cultural, legal, political and/or organizational. Reverse salients in CI projects may include metadata, intellectual property rights, federating databases with different equipment and formats, reluctance of modelers to switch to a better program if the learning curve is too steep, domain-specific data-sharing and publication cultures, lack of incentives in universities for CI-building and data sharing work, and the inability to translate salients across different fields.

Gateways. Gateways are the "plugs and sockets that allow new networks to be joined with existing networks with minimal constraint."[311] They are usually the joining together of a technical solution and a social choice, such as standards, that are integrated into users' communities of practice. Gateways work best without specifying exactly how work is to be done, or how information is to be processed.[312] Email, for example, is an extremely successful gateway because it can be used for almost any task.

Path Dependence. Path dependence makes infrastructures very difficult to alter once they are established. It occurs when early choices can constrain available options later on in the development process. Thus, decisions at the beginning of the development process can have an important effect on the long-term development of the infrastructure.

Key components of path dependence are localized learning, irreversibility, and network effects. *Localized learning* refers to choosing between satisfaction and optimization, and adapting to investments rather than considering alternatives. *Irreversibility* refers to the situation when infrastructure alternatives become too costly to consider due to the amount of money, time, attention, retraining and coordination required to change course. *Network effects* refer to technologies whose value increases exponentially with widespread adoption (the telephone, for example).

There are positive and negative elements of path dependence. One positive aspect is that progress is built on existing infrastructure. Negative aspects include inefficiencies, inheriting and entrenching defects of existing infrastructure, and integration of ad hoc systems into larger networks, which can lead to bad choices.

Scale Effects. "Cyberinfrastructure developers are focused on transforming small-scale, short-term, local projects into large-scale, functional infrastructures. The pattern of history suggests that this will take a long time — on the order of decades."[313] These projects typically follow an 'S-curve' pattern, usually lasting 40-50 years, in which an initial period of linear growth is followed by exponential growth and then a slow period of linear growth as the infrastructure reaches its maximum extent.

TENSIONS

[309] Edwards et al., "Understanding Infrastructure," 14-19.
[310] Ibid., 14.
[311] Ibid., 16.
[312] Paul Forster and John Leslie King, "Information Infrastructure Standards in Heterogeneous Sectors: Lessons from the Worldwide Air Cargo Community," in Kahin, Brian and Janet Abbate, eds., *Standards Policy for Information Infrastructure,* (Cambridge, MA: MIT Press) 148-77, quoted in Edwards et al., "Understanding Infrastructure," 17.
[313] Edwards et al., "Understanding Infrastructure," 19.

Infrastructure development often faces a series of tensions that challenge the idea that infrastructure building is planned, orderly, or mechanical.[314] IT is necessary to resolve social, cultural, and organizational issues in tandem with IT services in order to create a strong and lasting infrastructure. Tensions identified by Edwards et al. include: 1) time, scale and agency; 2) winners and losers; 3) interest and exclusion; 4) ownership and investment models; 5) data cultures and data tensions; 6) data storage and meaning; 7) data sharing; and 8) trust.[315] The long-term success of any CI project is dependent on long-term, focused attention to these issues.

Time, Scale, and Agency. Time refers to short-term funding decisions versus the long-term time-scale needed for infrastructures to grow and become established (usually many decades). Scale refers to choices between worldwide interoperability and local optimization. Agency encompasses how to navigate planned versus emergent change.

Winners and Losers. "Winners and losers" refers to those who benefit from, and those who are left behind, as a result of certain decisions or the course of infrastructure development. This process includes," a good deal of what economists have labeled 'creative destruction' as practices, organizations, norms, expectations, and individual biographies and career trajectories bend – or don't – to accommodate, take advantage of, and in some cases simply survive the new possibilities and challenges posed by infrastructure."[316]

Interest and Exclusion. "Interest and exclusion" refers to the unequal distribution of power and consequences of infrastructural change among stakeholders and contributors. It can occur when as a conflict between "design-centric" and "user-centric" visions of infrastructure development. In addition, established stakeholders and developers, who owe their status and power to the current infrastructure, are unwilling to change and thus marginalize other actors.

Ownership and Investment. Tensions of ownership and investment occur when changing infrastructures run into intellectual property rights, public/private investment models, and policy objectives. Important questions of ownership include:[317]

1. *"Who "owns" the results of deeply collaborative work?*
2. *By what mechanisms can (or should?) downstream revenues from such work be distributed?*
3. *How far does (or should?) property in "raw" data extend when the reworking of community repositories leads to new results?*
4. *Where researchers work against a mixed background of publicly owned and privately held intellectual property, how are rights to use and compensation to be balanced?"*

Tensions around data ownership and investment are likely to grow with more collaborative forms of research. Questions of ownership may arise when academic research comes in contact with industry-based research, when national policy objectives meet the global pull of science, or when data handling, sharing and collaboration for new CI projects challenge existing intellectual property rights.

Data Cultures and Data Tensions. Data storage, preservation, and curation are at the center of some of the most complex challenges to CI development. These tensions are best described in this way: "In the daily working world of science, infrastructural tensions and conflicts are very often played out and resolved (or not) at the level of data. Data, and the anxieties and tensions it occasions, represents the front

[314] Edwards et al., "Understanding Infrastructure," i.
[315] Ibid., 24-33
[316] Ibid, 24
[317] Ibid, 30

line of cyberinfrastructure development…and…in some ways…the target of its highest ambitions. From one view at least, cyberinfrastructure is principally about data: how to get it, how to share it, how to store it, and how to leverage it into the major downstream products (knowledge, discoveries, learning, applications, etc.) we want our sciences to produce."[318] Data have multiple meanings, and there is much variation in meaning across different domains. Data may encompass things (samples, specimens, etc.), or it may be the result of a model, or the model itself. These differences raise important questions regarding data storage, preservation, curation, and re-use.

Data Storage and Meaning. Data storage and meaning refers to the storage, preservation and curation of data. A key question is how much data, in what form, is needed to accomplish what task? The answer depends on the short and long-term data audience, and purpose of the data itself. Other issues to consider are data re-use (determining how much metadata is needed to support future, alternative, comparative use, for example), and the fact that data have different meanings across science domains.

Data Sharing. Data sharing is also a common tension in CI development, and necessitates tools that have the ability to accommodate and translate different needs from distinct domain communities. These tools must allow people from one community to make sense of data generated from another community. Technical solutions can help solve problems of data sharing but do not answer underlying questions of incentives, organization, trust and local variation. For example, some domains have advanced, highly-structured data-sharing norms, while others have no defined norms and procedures for data-sharing.

Trust. Data trust is a key component of data sharing. Data trust issues involve data generators and data users and can be summarized by the following questions:

- *"Can I trust those I agree to share my data with to make reasonable and appropriate use of it, and on a timeline which doesn't impede my own requirements re: publication, credit, and priority?"[319]*
- *"On the receiving end, can I trust the data I'm getting, particularly as collaborative webs lengthen and my first-hand knowledge of the data and its producers recedes?"[320]*

Questions of trust require social, organizational, cultural and technical solutions. Edwards et al. argue that they must be resolved in order for data integration to succeed; "Where uncertainty exists, and where two data cultures collide, well-intentioned efforts to promote sharing via technical or organizational fixes are unlikely to succeed. This aspect of scientific work, absolutely central to the daily practice of scientists, is arguably the one least well understood within the current cyberinfrastructure development process." [321]

CYBERINFRASTRUCTURE DESIGN RECOMMENDATIONS

Edwards et al. offer a number of CI design recommendations[322] that address the dynamics and tensions described above.

Purpose. CI design should permit collaborative work that enables scientists from different domains to work together. Changes in how scientific work is organized should be built into the CI design to achieve greater interdisciplinary collaboration. For example, reward structures can be constructed to encourage database work or early sharing of results.

Boundary Work. CI can bridge the 'great divide' between system building and social analysis.[323] Edwards et al. recommend the Scandinavian Community Supported Cooperative Work (CSCW) community to

---

[318] Edwards et al., "Understanding Infrastructure,"31
[319]Ibid, 32
[320] Ibid, 32
[321] Ibid, 33
[322] Edwards et al., "Understanding Infrastructure," 33-37.

encourage collaborative work. In the CSCW model, ethnographers are trained to act as brokers between IT developers and users because Scandinavian law requires that the end-user community co-design IT to be introduced into the workplace. Edwards et al. also recommend an analysis of social networks to chart communication patterns in distributed organizations and incorporate lessons learned from past CI projects.

Principles of Navigation. Edwards et al. argue that there is not a specific way to design CI; therefore, it is necessary to establish a set of design principles of navigation instead of a rigid roadmap. Because it can be difficult to determine whether something needs a technical or organizational solution, a set of flexible navigation principles that recognize and respond to path dependence and unanticipated barriers to adoption can greatly enhance the likelihood of CI success.

Standards and Flexibility. Flexibility should be included in CI design to avoid path dependence and facilitate change. The authors argue, "To design for flexibility, at each stage, with each dynamic, and in every type of intervention or evolution, possibilities for seeking alternative pathways must always be kept open."[324] Standards should be employed as a means to create flexibility, although this requires a different mindset regarding standards. Gateways should be flexible and compatible among political, operational and technical domains.

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE CI PROJECTS

Edwards et al. argue that it is important to learn from past successes and failures, noting that "infrastructural development is always a contested process, tied as it is to questions around access, power, and the life chances of groups and individuals."[325] To improve future CI, they suggest learning from past CI, improving CI analysis, and enhancing CI resiliency and research.[326]

A comparative social scientific analysis of CI projects, use of accurate and realistic reporting mechanisms, aligning extant or creating new incentive structures, and "instrumenting" CI projects for social scientific analysis will promote learning from past CI projects. Infrastructural diagnostics, training for information managers, cross-disciplinary symposia for early-career scientists, and fitting funding to CI-development time scales can improve current CI analysis. CI resiliency, sustainability and reach can be enhanced through design flexibility, mechanisms to incorporate under-represented groups and institutions, and alliances with niche organizations to eliminate redundancies in expertise.

HISTORICAL CASE STUDIES

PROJECT MOHOLE

Project Mohole (1957-1966) provides an excellent example of the importance of a governance framework and the necessity to address the dynamics, tensions, and design issues described by Edwards et al. Project Mohole ultimately failed because it had no formal governance framework and was not able to recognize or resolve many of the issues that plagued the project from the beginning, including a lack of clear purpose and objectives, leadership and a higher authority to resolve disputes, establish decision-making processes, launch communication mechanisms, and clearly assign responsibility and accountability. Instead of putting in place a governance framework and trying to resolve some of these issues before the project began, project developers and

[323] Geoffrey Bowker, Susan Star, Les Gasser, and William Turner, (ed.), *Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide*. (Mahwah, NJ: Erlbaum, 1997), quoted in Edwards et al., "Understanding Infrastructure," 34.
[324] Edwards et al., "Understanding Infrastructure," 36
[325] Ibid., 33
[326] Ibid., ii.

stakeholders charged full speed ahead, thus leading to one of the most expensive, and most catastrophic, failures of a big research project in modern history. The summary below is based on DS Greenburg's "Mohole: The Project that Went Awry" (1964, 1966).[327]

In 1957, the Earth Sciences Review Panel of the National Science Foundation decided to obtain a sample of the Mohorovičić discontinuity or "Moho," which lies 3-6 miles below the ocean floor. It was not clearly stated whether the objective was just to drill to the Moho or to conduct a far-reaching, multilevel drilling program. It was also unclear as to who was ultimately in charge of and accountable for the project, who had the authority to resolve disputes, and who was responsible for creating and maintaining project objectives.

Despite these problems, the proposal received funding, and the first phase of Project Mohole began. Phase I consisted of record-setting test drillings in the Pacific Ocean, and it was completed in a few weeks, staying within its budget. Thus, Phase 1 was considered to be a great success.

Phase II of Project Mohole was a disaster. At this time, the technology to drill to the Moho did not exist, and the governance mechanisms to resolve the social, cultural, technological, and scientific barriers to achieve Project Mohole's goals were not in place. Although a contractor was hired to design and implement the technology to carry out Phase II drilling, many important questions remained unresolved, including: What were the ultimate objectives of the project? What type of ship and drilling platform should built? Who had the authority to make that decision?

The project was stalled by disputes between the NSF, the contractor, the oceanographers, Congress, and other stakeholders. The controversy was so great that the Bureau of Budget froze funding in 1964, and Project Mohole was finally discontinued in 1966 when the Senate cancelled funding. The lack of clear purpose and clear organizational structure had caused Project Mohole to become a political liability, a money pit, and an example of how not to run a big research program.

LEADERSHIP PRINCIPLES IN NSFNET, UNIDATA, DLESE, AND NSDL

"Collaboration, Alignment and Leadership" by David Fulker (2008)[328] employs leadership principles to compare and contrast four NSF projects that required a high level of collaboration and a collective sense of identity and mission: National Science Foundation Network (NSFNET), Unidata, Digital Library for Earth System Science (DLESE), and the National Digital Science Library (NSDL). Each of these projects had very different outcomes. NSFNET has expanded into the Internet as we know it and received little to no NSF support in 2008. Unidata experienced a gradual expansion and is an important part of the NSF/Atmosphere directorate. DLSESE stopped receiving funding from the NSF/Geoscience directorate and in 2008.

The leadership principles upon which Fulker bases his analysis are:

1. **Disciplined Thought:**[329] Willingness to confront reality; A simple, coherent strategic concept.
2. **Disciplined Action:**[330] Willingness to say "no;" Clear identification of who holds responsibilities; Advancement as a cumulative process.
3. **Persistent Core:**[331] An enduring purpose; Immutable core values (around which to advance).

[327] See Greenberg, DS, "Mohole: The Project That Went Awry." *Science* 143, no. 3602 (1964): 115-119, Greenberg, DS, "Mohole: The Project That Went Awry (ii)." *Science* 143, no. 3603 (1964): 223-227, Greenberg, DS, "Mohole: The Project That Went Awry (iii)." *Science* 143, no. 3604 (1964): 334-33, and Greenberg, DS, "NSF Appropriation: Mutiny on the Mohole." *Science* 152, no. 3724 (1966): 895-896

[328] David Fulker, "Collaboration, Alignment and Leadership, invited piece in 'NDSL Reflections," an online series accessed at http://nsdlreflections.wordpress.com, 2008.

[329] Jim Collins, *Good to Great: Why Some Companies Make the Leap and Others Don't*, (New York: NY, HarperCollins), 2001, 65-199, quoted in David Fulker, "Collaboration, Alignment and Leadership", 4.

[330] Jim Collins, *Good to Great*, 2001, 123-143, quoted in David Fulker, "Collaboration, Alignment and Leadership", 4.

[331] Jim Collins, *Good to Great*, 188-201, quoted in David Fulker, "Collaboration, Alignment and Leadership", 4.

4. **Primary alignment with one of three value disciplines** (while maintaining the other two):[332] Operational excellence; Customer relations; Product Leadership.

Fulker argues that NSFNET and Unidata were better able to confront the realities that most new initiatives face, while NSDL and DLESE were less able to agree on how to limit their expectations, and were not able to solidify a common sense of purpose and identity. NSDL and DLESE valued diversity over coherence, leading to "imprecise alignments among the collaborators' visions and priorities,"[333] in addition to difficulties in gaining collaborator agreements on common values and target audiences.

In terms of willingness to say "no," NSFNET and Unidata's grant programs gave better clarification and support for leadership decisions about what not to do. DLESE's and NSDL's grant programs did not provide good support for leadership decisions about what not to do, so they had trouble making decisions regarding focus. NSFNET and Unidata were able to maintain the leadership principle of advancement as a cumulative process because they received steady support from NSF. This support was essential because it took Unidata 10 years to grow a robust community. In addition, NSFNET and Unidata's original purposes withstood the test of time, while NSDL and DLESE had difficulties articulating clear statements of purpose.

Fulker found that all four projects had immutable core values. He found little variation among the projects regarding this leadership component because each program was built around the core values shared by stakeholders.

Fulker found much greater variation regarding primary alignment with a single value discipline (see Fulker's #4 above). NSFNET focused on operational excellence, while Unidata started with customer relations and then shifted to product leadership. In contrast, Fulker argues that community leaders and NSF contributed to NSDL and DLESE ambiguities in regards to the three value disciplines discussed above. Because NSDL, and DLESE had not established a clear concept, purpose, and target constituents, they were unable to achieve alignment to a single value discipline.

Fulker also highlights the role of the community in leading to leadership and purpose ambiguities in NSDL and DLESE. Although there was general agreement for a Science, Technology, Engineering, and Math (STEM) library, there was a lack of agreement on specific needs and how to meet these needs, which led to difficulties in community-building. In NSDL, more developers were involved than users, which led to a lack of focus on end-user needs. Thus, in the case of NSDL, "build it and they will come" did not work.

To alleviate the problems experienced by NSDL and DLESE, Fulker recommends the creation of an evaluation framework with the involvement of experts such as Jim Collins, Michael Treacy and Fred Wiersema, authors of the four points highlighted by Fulker. An evaluation framework would eliminate redundancies and explain principles in the specific terms of the evaluation framework, thus avoiding informal and reflective analysis.

CONCLUSION: SECTION 2

HISTORICAL INFRASTRUCTURE DEVELOPMENT

Common Steps:

System building: Systems are linked into networks, leading to distributed control and access.

[332] Michael Treacy and Fred Wiersema, "Customer Intimacy and Other Value Disciplines," *Harvard Business Review*, 71 (1993): 84-93, quoted in David Fulker, "Collaboration, Alignment and Leadership", 4.
[333] David Fulker, "Collaboration, Alignment and Leadership," 5.

Technology Transfer: Technology is transferred from one location or domain to another and the system adapts to new conditions.
Integration and Consolidation: Heterogeneous systems are linked into networks through gateways, consolidating the infrastructure project.

Infrastructure Dynamics

Reverse Salients: Reverse salients are critical, unsolved problems, defined as, "points where an advancing front is held back," and can be technical, social, cultural, legal, political and/or organizational.
Gateways. Gateways are the "plugs and sockets that allow new networks to be joined with existing networks with minimal constraint." They are usually the joining together of a technical solution and a social choice, such as standards, that are integrated into users' communities of practice.
Path Dependence. Path dependence makes infrastructures very difficult to alter once they are established. It occurs when early choices can constrain available options later on in the development process.
Scale Effects: It takes 40-50 years for short-term, local infrastructure projects to grow into large-scale, functional infrastructures.

Infrastructure Tensions

Time, Scale, and Agency. Time (short-term funding decisions versus the long-term time-scale needed for infrastructures to grow); Scale (choices between worldwide interoperability and local optimization); Agency (how to navigate planned versus emergent change).
Winners and Losers. Those who benefit from, and those who are left behind, as a result of certain decisions or the course of infrastructure development.
Interest and Exclusion. The unequal distribution of power and consequences of infrastructural change among stakeholders and contributors.
Ownership and Investment. Tensions of ownership and investment occur when changing infrastructures run into intellectual property rights, public/private investment models, and policy objectives.
Data Cultures and Data Tensions. Data storage, preservation, and curation are at the center of some of the most complex challenges to CI development.
Data Storage and Meaning. Data storage and meaning refers to the storage, preservation and curation of data. A key question is how much data, in what form, is needed to accomplish what task?
Data Sharing. Data sharing necessitates tools that have the ability to accommodate and translate different needs from distinct domain communities.
Trust. Data trust is a key component of data sharing and involves data generators and data users.

CI Design Recommendations

Purpose. CI design should permit collaborative work that enables scientists from different domains to work together.
Boundary Work. CI can bridge the 'great divide' between system building and social analysis.
Principles of Navigation. It is necessary to establish a set of design principles of navigation instead of a rigid roadmap because there is no one way to design CI.
Standards and Flexibility. Flexibility should be included in CI design to avoid path dependence and facilitate change.

Recommendations for Future CI Projects

Learn from successes and failures of past CI projects. A comparative social scientific analysis of CI projects, use of accurate and realistic reporting mechanisms, aligning extant or creating new incentive structures, and 'instrumenting' CI projects for social scientific analysis will promote learning from past CI projects.
Improve current CI analysis. Infrastructural diagnostics, training for information managers, cross-disciplinary symposia for early-career scientists, and fitting funding to CI-development time scales can improve current CI analysis.

Enhance CI resiliency and research. CI resiliency, sustainability, and reach can be enhanced through design flexibility, mechanisms to incorporate under-represented groups and institutions, and alliances with niche organizations to eliminate redundancies in expertise.

HISTORICAL CASE STUDIES: PROJECT MOHOLE

Lack of governance structure led to failure of project. Specific, unresolved issues that contributed to project failure included lack of clear purpose and objectives, leadership and a higher authority to resolve disputes, decision-making processes, communication mechanisms, and clear assignment of responsibility and accountability. A governance framework could have created mechanisms to resolve these issues. A governance framework needs to be put in place from the very beginning of a project in order to deal with anticipated and unanticipated challenges and issues as the occur.

LEADERSHIP PRINCIPLES IN NSFNET, UNIDATA, DLESE, NSDL

NSFNET, Unidata, DLESE, and NSDL were successful (or not) based on the degree of clarity of leadership, purpose, accountability, and primary alignment to single value discipline (see # 4 in Key Leadership Principles below). Each project had a different outcome:

NSFNET Expanded into the Internet as we know it and received little to no NSF support in 2008.
Unidata experienced a gradual expansion and is an important part of the NSF/Atmosphere directorate.
DLSESE stopped receiving funding from the NSF/Geoscience directorate.
NSDL's 'core integration' component was being replaced by grants as of 2008.

Key Leadership Principles

Disciplined Thought: Willingness to confront reality; A simple, coherent strategic concept.
Disciplined Action: Willingness to say "no;" Clear identification of who holds responsibilities; Advancement as a cumulative process.
Persistent Core: An enduring purpose; Immutable core values (around which to advance).
Primary alignment with one of three value disciplines (while maintaining the other two): Operational excellence; Customer relations; Product Leadership.

Leadership Recommendation

Create an **evaluation framework** with the involvement of leadership experts. An evaluation framework would eliminate redundancies and explain principles in the specific terms of the evaluation framework, thus avoiding informal and reflective analysis.

SECTION 3: CURRENT STATE OF GOVERNANCE

This section moves from historical infrastructure and cyberinfrastructure creation and case studies to more contemporary issues in governance. Several governance case studies provide examples of governance framework in existing domain science (such as IRIS, Unidata, UCAR, CUAHSI, IEDA), IT (such as W3C, XSEDE, Apache, and Slashdot), federated (such as ESIP, OGC, DataONE), and large facilities groups (such as OOI, EarthScope, NEON). In addition, recommendations from recent geoinformatics workshops and government CI reports are also reviewed. These case studies and reports clearly identify that many of the issues faced by today's EarthCube community are not new and that the community has been converging on a path forward in cyberinfrastructure for the geosciences for close to a decade.

GOVERNANCE CASE STUDIES

For each organization we provide 1) a general background of the organizations mission and membership (if applicable), 2) the funding mechanisms in place, 3) the current governance framework, and 4) staffing of the organization (either paid, volunteer, or a combination).

Case studies include:

- Domain science; CUAHSI, IRIS, Unidata
- IT: W3C Consortium, DataNet Federation Consortium, DataONE, Apache Software Foundation
- Federations: ESIP and OGC
- Large Facilities: SURA US IOOS Coastal and Ocean Modeling Testbed, Project GENI

DOMAIN SCIENCE GROUPS: CUAHSI, IRIS, UNIDATA

CONSORTIUM OF UNIVERSITIES FOR THE ADVANCEMENT OF HYDROLOGIC SCIENCE (CUAHSI)

General Background. The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) represents over 130 U.S. universities and international hydrologic-science organizations. Its mission is to enable "the university water science community to advance the understanding of the central role of water to life, Earth, and society."[334] It is composed of hydrologic databases. CUAHSI supports its community to "advance water science and to improve societal well-being by:

- Developing, supporting, and operating research infrastructure;
- Improving and promoting access to data, information and models;
- Articulating and advocating priorities for community level water-related research and observations
- Facilitating interactions among the diverse water research community
- Promoting interdisciplinary education centered in water science
- Translating scientific advancements into effective tools for water management and policy."[335]

Funding. The CUAHSI Hydrologic Information System (HIS) is an NSF-funded project whose mission is to improve access to water data through the development of infrastructure and services.[336] Specific services include tools, standards and procedures, and Web services. The CUAHSI-HIS architecture is composed of three main components: data publication, data cataloging, and data discovery, access, and analysis.

---

[334] "What is CUAHSI?" Consortium of Universities for the Advancement of Hydrologic Science, Inc., accessed June 1, 2012, http://www.cuahsi.org/docs/What-is-CUAHSI.pdf.
[335] Ibid.
[336] Ibid.

Governance. The CUAHSI governance framework is guided by a set of by-laws:[337]

- **Article I: Name**
- **Article II: Member Institutions:** Describes how institutions may become CUAHSI Members or Affiliates.
- **Article III: Meetings of Membership and Election of the Board of Directors:** Provides requirements for the annual meeting, notice of meetings, quorum, nominating committee, ballot, election of directors, method of voting, counting of ballots, special elections, special meetings, and virtual participation.
- **Article IV: Board of Directors:** Describes the powers, composition, election, term of office and processes for resignation, removal, transfers and vacancies of Board Members.
- **Article V: Meetings of the Board of Directors:** Provides requirements for the annual meeting, special meetings, notice of meetings, quorum, voting, action without a meeting, and virtual participation.
- **Article VI: Officers:** Describes officer qualifications, roles of the Chair Person, President, Secretary, and Treasurer, elections and term of office, and the processes for resignation, removal and vacancies of Officers.
- **Article VII: Executive Committee of the Board, Other Committees and Advisory Council:** Describes the process for the creation of new committees, the powers of the Executive Committee, the role of the Internal Audit Committee, Special Committees, Standing Committees, Senior Advisory Council, and the Corporate Advisory Board.
- **Article VIII: Election of Officers and Executive Committee at the Annual Meeting of the Board of Directors:** Describes the nominating and voting processes to elect officers and Executive Committee members at the annual meeting.
- **Article IX: Fees and Assessments:** Requires that new members contribute an initiation fee and additional fees as determined by majority vote of the Board of Directors.
- **Article X: Compensation**
- **Article XI: Indemnification**
- **Article XII: Fiscal Year**
- **Article XIII: Seal of the Corporation**
- **Article XIV: Amendments to the Bylaws:** The CUAHSI bylaws may be amended by an affirmative vote of 60% of the Members, Non-profit Affiliates and International Affiliates.

Staff. CUAHSI is governed by the Board of Directors and managed by the Executive Committee (composed of Board members and Officers). CUAHSI Committees and a small paid staff carry out CUAHSI activities (Figure 20).

FIGURE 20: CUAHSI LEADERSHIP STRUCTURE[338]

INCORPORATED RESEARCH INSTITUTIONS FOR SEISMOLOGY (IRIS)

General Background. "IRIS [Incorporated Research Institutions for Seismology] is a consortium of over 100 US universities dedicated to the operation of science facilities for the acquisition, management, and distribution of

[337] See "Bylaws of the Consortium of Universities for the Advancement of Hydrologic Science, Inc.," cuahsi.org, adopted December 7, 2010, http://www.cuahsi.org/by-laws.html.
[338] "Organization," Consortium of Universities for the Advancement of Hydrologic Science, Inc., accessed June 1, 2012, http://www.cuahsi.org/organization.html.

seismological data. IRIS programs contribute to scholarly research, education, earthquake hazard mitigation, and verification of the Comprehensive Nuclear-Test-Ban Treaty."[339] In addition, IRIS functions as an interface between the scientific community, funding agencies and IRIS programs, and facilities are multi-user and multi-use.
The IRIS mission is to do the following:

- "Facilitate and conduct geophysical investigations of seismic sources and Earth properties using seismic and other geophysical methods.
- Promote exchange of geophysical data and knowledge, through use of standards for network operations, data formats, and exchange protocols, and through pursuing policies of free and unrestricted data access.
- Foster cooperation among IRIS members, affiliates, and other organizations in order to advance geophysical research and convey benefits from geophysical progress to all of humanity." [340]

Funding IRIS was originally funded with support from the National Science Foundation. IRIS is a 501(c)(3) nonprofit organization incorporated in the state of Delaware with its primary headquarters located in Washington, DC.

Governance. The IRIS governance framework establishes community ownership of instrumentation resources and data, and the framework is designed to allow IRIS to adapt to changes in technology and in the growth of the Consortium. IRIS's community governance model provides a framework for decision-making that preserves community oversight of IRIS facilities. IRIS is governed by Consortium members (114 US research universities and research institutions – see Figure 21 below).

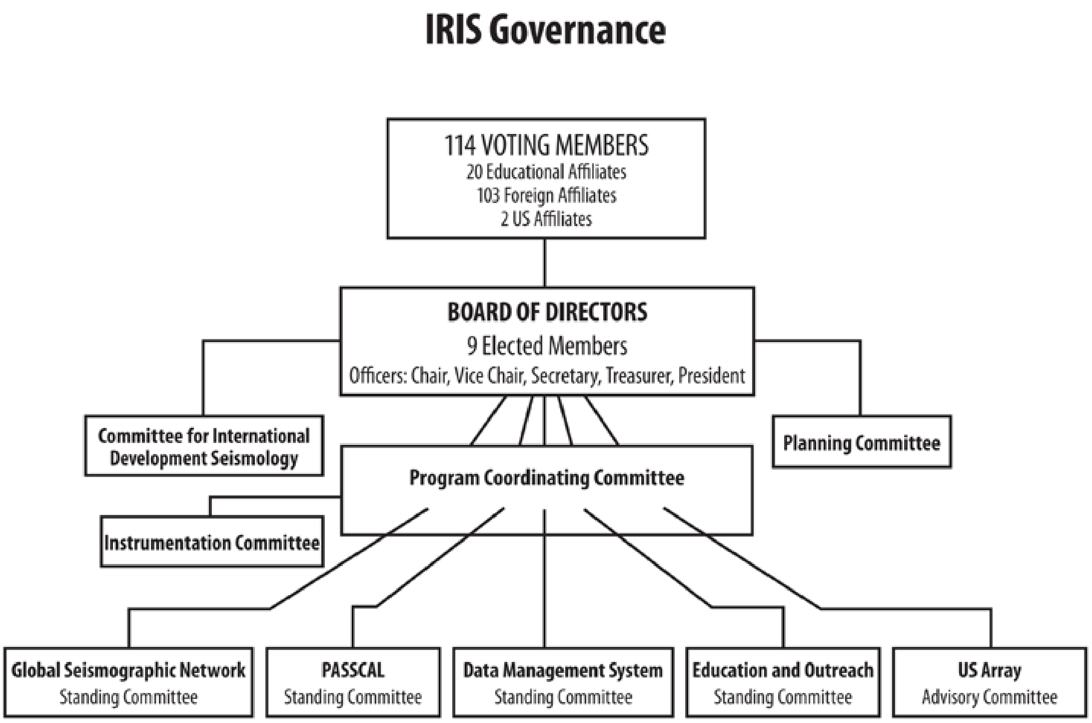

FIGURE 21: IRIS GOVERNANCE[341]

Standing Committees (USArray Advisory, Planning, Program Coordination, Instrumentation) and Board of Directors and Subcommittees make decisions regarding planning and operations. In addition, IRIS employs workshops, annual meetings, symposia, and newsletters to promote on-going interactions with scientists and

[339] IRIS, "About IRIS," last modified 2012, http://www.iris.edu/hq/about_iris.
[340] Ibid.
[341] IRIS, "About IRIS," last modified 2012, http://www.iris.edu/hq/about_iris.

member institutions and to receive feedback. These mechanisms allow the research community to communicate its evolving needs to funding agencies.

UNIDATA

General Background. Unidata is a scalable, loosely-federated model of more than 250 institutions brought together under the common goal of sharing data and tools to access and visualize data. Unidata is funded primarily by NSF and is one of eight programs in the University Corporation for Atmospheric Research (UCAR) Office of Programs (UOP). Unidata is composed of a National Center and Discipline-Specific Centers that interact directly and support their communities. The Unidata Program Center: [342]

- "Explores new technologies
- Evaluates and implements technological standards and tools
- Advocates for the community
- Provides leadership in solving community problems in new and creative ways
- Negotiates for new and valuable data sources
- Facilitates data discovery and use of digital libraries
- Enables student-centered learning in the Earth system sciences by promoting use of data and tools in education
- Values open standards, interoperability, and open-source approaches
- Develops innovative solutions and new capabilities to solve community needs
- Stays abreast of computing trends as they pertain to advancing research and education"

Funding. The Unidata program is staffed with funding from NSF/AGS (core funding) and other grants from NSF, NOAA, NASA, and other organizations.

Governance. Unidata is governed by a Policy Committee and a User's Committee (standing committees), in addition to an Implementation Working Group (in the early days of the program) and Steering Committees formed on an as-needed basis. The Policy Committee advises on the direction and program priorities in addition to making sure university community needs are met. The Users Committee seeks user input and feedback.

The Unidata governance framework consists of representation by community members who are already active in their communities; tight feedback loops on program projects; direct lines of communication among the community, governing bodies, and staff; and a consensus-oriented and pragmatic approach to decision-making. Governance principles include the following:[343]

- Community-led, autonomous, accountable to community needs
- Open, transparent, inclusive
- Active, independent, informed oversight
- Consensus-oriented
- Balanced representation
- Shared responsibility
- Direct lines of communication
- Focus on long-term strategic direction

Unidata employs a number of governance mechanisms to put its governance framework into practice. These mechanisms include active, independent, and informed oversight and stewardship, consensus-oriented and

[342] "Overview of Unidata," Unidata, accessed June 1, 2012, http://www.unidata.ucar.edu/about/#overview.
[343] Ramamurthy, Unidata Governance: A Quarter Century of Experience, 4-5

pragmatic approaches to decision making, balanced representation from the full range of participating institutions, shared responsibility between the governing committees and the program with clear delineation of the roles, direct lines of communication between governing committees and staff, and focus on long-term strategic direction with nimbleness to respond to rapid changes. Unidata Governance challenges include how to best scale efforts to meet the needs of growing communities, how to balance the need to develop new solutions with the need to provide support for existing solutions, and how to do all these things with limited resources.
Staff. The Unidata Governing Committee Members are all volunteers. However, the Unidata program, with funding in the core award, pays the travel expenses of Governing Committee Members to attend meetings but does not remunerate for their participation.

IT GROUPS: APACHE SOFTWARE, DATANET, DATAONE, W3C

APACHE SOFTWARE FOUNDATION

General Background. While there is not an official list, these six principles have been cited as the core philosophy behind the foundation, which is normally referred to as "The Apache Way":[344]

- collaborative software development
- commercial-friendly standard license
- consistently high quality software
- respectful, honest, technical-based interaction
- faithful implementation of standards
- security as a mandatory feature

Governance. Apache governance is meritocracy-based project promotion.

Project Management Committee (PMC) structure is defined as follows: "the role of the PMC is to further the long term development and health of the community as a whole, and to ensure that balanced and wide scale peer review and collaboration does happen. Within the ASF we worry about any community which centers around a few individuals who are working virtually uncontested. We believe that this is detrimental to quality, stability, and robustness of both code and long term social structures."[345]

DATANET FEDERATION CONSORTIUM

General Background. These comments are based on the DataNet Federation Consortium, an NSF funded project to build national data cyberinfrastructure for NSF funded research projects. Current participants include Ocean Observatories Initiative, Hydrology projects, and the iPlant Collaborative. Policies are ideally developed independently by each project for governance, management, administration, and assessment. The data grid (iRODS) allows each group to enforce their own policies, independently of the other projects. The infrastructure is designed to be highly extensible, separating policy enforcement from data manipulation, data storage, and data access mechanisms. While the software infrastructure is generic, the policies and procedures can be unique to each user community.

Funding. iRODS software development is supported under an NSF SDCI grant. Applications of the software include a DataNet Federation Consortium grant, projects with the iPlant Collaborative, the Ocean Observatories Initiative, a LifeTime Library student digital library at SILS, a collaboration with RENCI on E-iRODS development and more.

---

[344] "How the ASF Works," Apache Software Foundation, last modified 2012, https://www.apache.org/foundation/how-it-works.html.
[345] Ibid.

Governance. The iRODS data grid governance framework is mapped in Table 18 below.

Decision	IT Principles	IT Architecture	IT Infrastructure Strategies	Business Application Needs	IT Investment
Archetype					
Business Monarchy					
IT Monarchy					
Feudal	X	X		X	
Federal			X		X
Duopoly					
Anarchy					

TABLE 18: IRODS DATA GRID GOVERNANCE FRAMEWORK.[346]

From the perspective of the iRODS data grid, governance is accomplished through each collaborating group building a consensus for:

- Purpose of collaboration (shared collection, shared research initiative)
- Properties that the collaboration environment will maintain (completeness, integrity, standards, access)
- Policies that enforce the desired properties (computer actionable rules)
- Procedures that derive the desired properties (computer executable workflows)
- State information that tracks results of applying procedures
- Assessment criteria that verify desired properties have been enforced

iRODS data grid implements this approach using a distributed rule engine, a rule language, and workflows composed by chaining basic functions. At the moment, the scale is:

- 74 policy enforcement points (locations in the middleware where it is determined whether a policy needs to be enforced)
- 250 micro-services that encapsulate basic operations
- 209 system state variables to track outcomes (plus arbitrary number of user-defined status flags, descriptive metadata, provenance metadata)
- For collaboration environments, the policies need to represent a consensus across the participating groups

Staff. The number of persons in the DICE Center who work on iRODS is 9. The number of persons at RENCI who work on iRODS is 5. The total number of paid staff is 14. Some of these persons are part time.

The software is open source. There is an active international community who provide extensions to the software. The contributors include DistributedBIO, IN2P3 (France), University of Liverpool, Academia Sinica (Taiwan), ARCS (Australia), LSU Peashare, and more. Collectively across the international collaborators, 50 persons typically attend a yearly iRODS user group meeting; the number of users of the technology is much larger.[347]

[346] Reagan Moore, personal email, May 31, 2012.
[347] Reagan Moore, personal email, June 7, 2012.

General Background. Data Observation Network for Earth (DataONE) "is the foundation of new innovative environmental science through a distributed framework and sustainable cyberinfrastructure that meets the needs of science and society for open, persistent, robust, and secure access to well-described and easily discovered Earth observational data… DataONE will ensure the preservation, access, use and reuse of multi-scale, multi-discipline, and multi-national science data via three principle cyberinfrastucture elements and a broad education and outreach program."[348] The DataONE mission is to "enable new science and knowledge creation through universal access to data about life on earth and the environment that sustains it."[349] The DataONE vision is to "be commonly used by researchers, educators, and the public to better understand and conserve life on earth and the environment that sustains it."[350]

Governance. DataONE is organized into a Principle Investigator, Executive Director, Director for Development and Operations, DataONE Leadership Team, External Advisory Board, and Working Groups. DataONE governance and management structure is depicted below (Figure 22). DataONE communicates its progress, activities, and events through regular newsletters, frequent new items, and via Twitter, Facebook, and LinkedIn. There is also a DataONE Users Group (Dug) for more active DataONE members.

[348] DataONE, "What is DataONE?" accessed June 5, 2012, http://www.dataone.org/what-is-dataone.
[349] Ibid.
[350] Ibid.

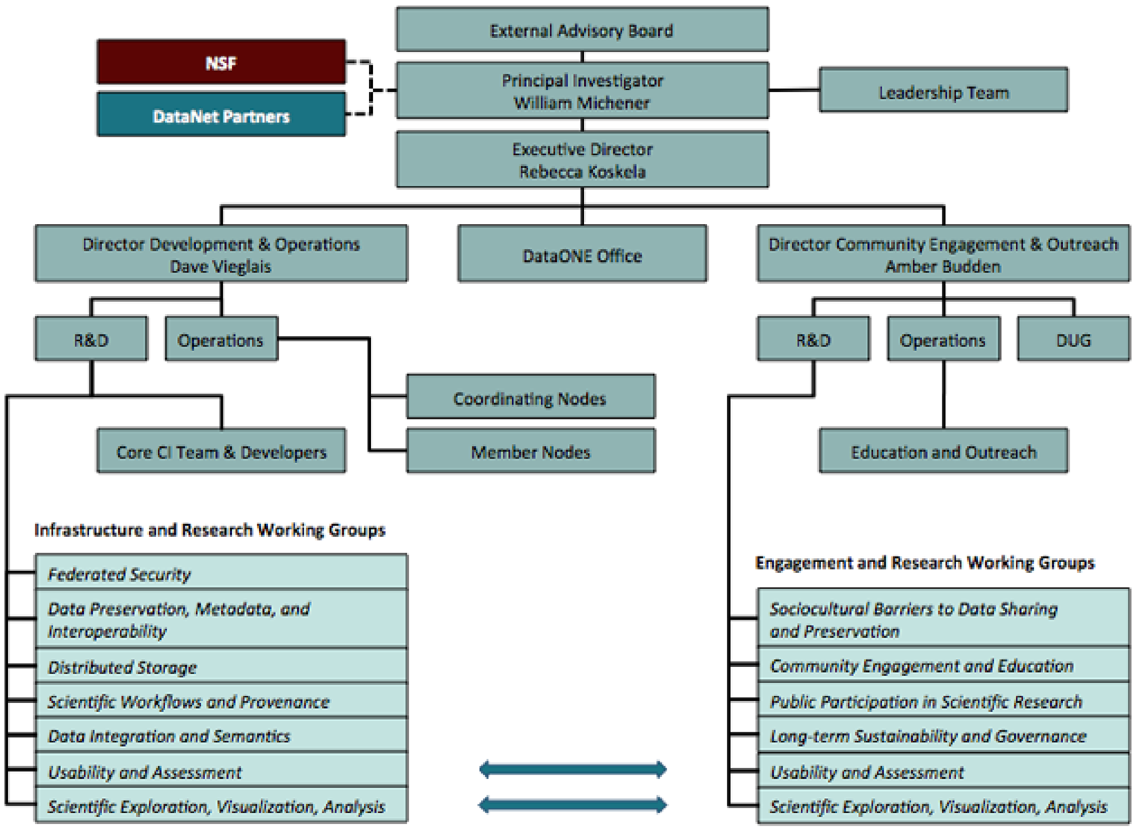

FIGURE 22: DATAONE ORGANIZATION[351]

DataONE community engagement and outreach includes working groups to identify, describe, and implement the DataONE cyberinfrastructure, governance, and sustainability models. Participatory, user-centered processes are employed for on-going evaluation of DataONE architecture. Specific goals of evaluation processes include: [352]

1. Identification and prioritization of stakeholder communities
2. Development of understanding of their perceptions, attitudes, and user requirements
3. Analysis and assessment of usability
4. Engagement of science teams in grand challenge science exemplars

Funding. "DataONE is a collaboration among many partner organizations, and is funded by the US National Science Foundation (NSF) under a Cooperative Agreement."[353]

W3C CONSORTIUM

General Background. The mission statement of the W3C Consortium is to "lead the World Wide Web to its full potential by developing protocols and guidelines that ensure the long-term growth of the Web. Below we discuss important aspects of this mission, all of which further W3C's vision of One Web."[354] W3C is carried out through a

[351] DataONE, "DataONE Organization," accessed June 5, 2012, http://www.dataone.org/organization.
[352] William Michener, Rebecca Koskela, Matthew Jones, Robert Cook, Mike Frame, Bruce Wilson, and David Vieglais, "DataONE-Enabling Cyberinfrastructure for the Biological, Environmental and Earth Sciences." National Science Foundation EarthCube White Paper: Designs Category, 2011.
[353] Ibid.
[354] W3C Consortium, "W3C Mission," last modified 2009, http://www.w3.org/Consortium/mission.

joint agreement among three "Host Institutions" (Massachusetts Institute of Technology - MIT, European Research Consortium for Informatics and Mathematics - ERCIM, and Keio University).

W3C membership is open to any commercial, educational, governmental entity, and individuals who want to join. W3C provides very specific details on potential Member roles and has a separate set of conditions describing how W3C applies its process to membership by projects. As of May 31, 2012 W3C has 368 members.[355] W3C's areas of focus include architecture, technology and society, and other domains (Figure 23).

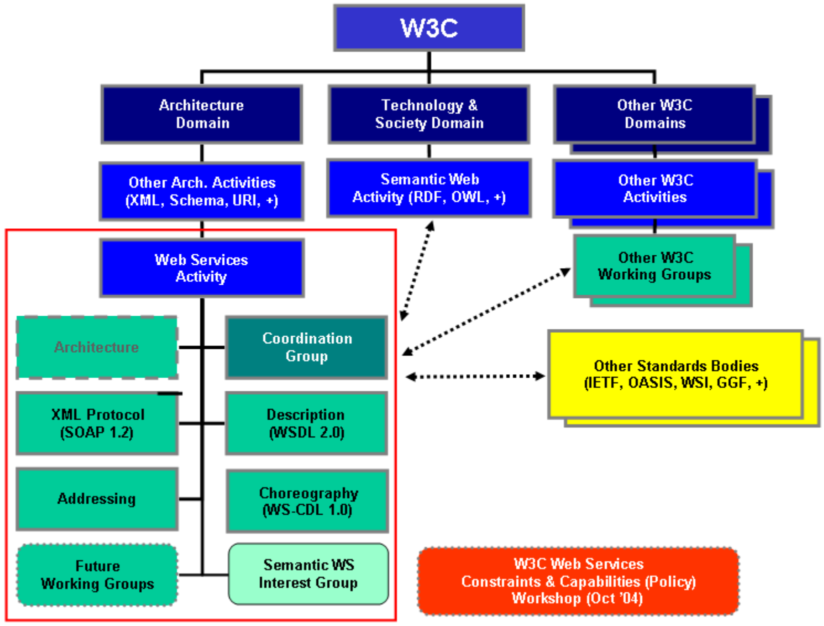

FIGURE 23: W3C ORGANIZATIONAL MODEL[356]

Funding. W3C sources of revenue include W3C Member dues, research grants and other sources of private and public funding, and sponsorship and donations.

Governance. The W3C governance framework is set up to carry out its mission statement and identifies a process for the standardization of Web technologies based on fairness, responsiveness, and progress. Standards development is based on community consensus, and all W3C stakeholders can provide input in the development of W3C standards. Several documents, including the W3C Process Document, Member Agreement, and Patent Policy, establish the roles and responsibilities of W3C consortium members. Key organizational components include:[357]

- **Advisory Committee:** each W3C member has a representative and participates in the W3C Process
- **Advisory Board:** members are elected by the Advisory Committee
- **Technical Architecture Group (TAG):** documents Web Architecture Principles
- **Chartered groups:** composed of Member representatives and invited experts and produces W3C deliverables

W3C fosters international participation in standards creation through a number of governance mechanisms. These mechanisms include liaisons with national, regional and international organizations; the Offices Program (promotion of W3C recommendations to developers, application builders, and other stakeholders); a policy for

[355] W3C, Consortium, "Current Members," last modified May 31, 2012, http://www.w3.org/Consortium/Member/List
[356] W3C Consortium, "Web Services Organization and Accomplishments," last modified 2004, http://www.w3.org/2004/Talks/1117-sb-gartnerWS/slide4-0.html.
[357] W3C Consortium, "Facts About W3C," last modified 2012, http://www.w3.org/Consortium/facts.

authorized translations of Web standards; forums around the world for people interested in standards creation; and W3C's Internationalization Activity, which helps to make sure the web is accessible to everyone.[358]
W3C protects the contributions of its members and developers through a Patent Policy for Working Groups. This document facilitates the development of W3C Recommendations by W3C Working Groups, promotes the implementation of Royalty-Free W3C Recommendations, and deals with patent issues that arise during and after the development of a Recommendation.[359]

Staff: A Director and CEO guide staff at W3C, and a small management team and staff of technical experts are in charge of strategic planning and allocation of resources.

FEDERATED GROUPS: ESIP, OGC

FEDERATION OF EARTH SCIENCE INFORMATION PARTNERS (ESIP)

Background information. [360] The Federation of Earth Science Information Partners (ESIP) is a broad-based, distributed community of science, data and information technology practitioners that coordinate interoperability efforts across the earth and environmental science communities. Participation in the ESIP Federation provides individual members an intellectual commons to expose, gather and enhance their own in-house capabilities in support of their organization's mandate. By virtue of working in the community, ESIP members experience the network effect, which enables more coordinated interoperability efforts across domain-specific communities. The ESIP Federation has a 14-year track record of success and continued growth using this community-based, discipline and agency neutral approach. These efforts catalyze connections across organizations, people, systems and data allowing for improved interoperability in distributed systems. The ESIP Federation is working in areas at the leading edge of technology development including data architecture, data management and preservation, tool creation, cloud computing, semantic web, data systems integration, and data management professional development.

Initially created by NASA in 1998, the ESIP Federation was formed in response to a National Research Council recommendation calling for the involvement of community stakeholders in the development of NASA's EOSDIS as a critical element of the U.S. Global Change Research Program (http://www.gcrio.org/USGCRP/LaJolla/cover.html). Since its inception, the ESIP Federation has continually grown and attracted a diverse group of partners, now including more than 135 member organizations. The ESIP Federation's membership is broad, including federal data centers, government research laboratories, research universities, education resource providers, technology developers, and various nonprofit and commercial enterprises. This diversity (multi-sector, agency, and function) provides expert capabilities on which participating organizations can draw. NASA and NOAA are the ESIP Federation's principal sponsors.

Funding. The ESIP Federation is principally supported by NASA and NOAA.

Governance Structure. The ESIP Federation is governed by its Assembly, the standing body which affords each partner one vote. Governance documents include a Constitution and Bylaws as well as a 5-year strategic plan (http://wiki.esipfed.org/index.php/Federation_Documents). In between annual Assembly business meetings, the ESIP Federation is governed by its Executive Committee; a body comprised of the ESIP Federation's elected President and Vice President, its Standing and Administrative Committee chairs, and members representing each ESIP type (Data Center, Research, and Application Developer). All Executive Committee members are either elected by the Assembly or by their ESIP Type caucus. While the ESIP Federation remains unincorporated, the Foundation

[358] Ibid.
[359] W3C, "Facts About W3C," last modified 2012, http://www.w3.org/Consortium/facts
[360] Carol Meyer, personal email, June 5, 2012.

for Earth Science (a 501(c)3 provides financial and other management services to the ESIP Federation.

Staff. The ESIP Federation runs with support from two professional staff and a host of volunteer contributions from experts across the earth science data and technology community.

OPEN GEOSPATIAL CONSORTIUM (OGC)

General Background. OGC is a Voluntary Consensus Standards Organization founded 1994. It has 445 members, 38 adopted standards, hundreds of product implementations, broad user community implementation worldwide, and has alliance partnerships with 30+ standards and professional organizations. OGC is member-driven. OGC's Approach for Advancing Interoperability includes the Interoperability Program (IP), Standards Program, Compliance Program, and Marketing and Communications Program.

Funding. OGC funding comes from membership dues and project fees (Interoperability Program initiatives).[361]

Governance. OGC's Interoperability Program, Standards Program, Compliance & Interoperability Testing & Evaluation (CITE) Program, and Marketing and Communications are governed by a Board of Directors, Strategic Member Advisory Committee, Global Advisory Council, and Executive Director (see figure 2 below). Each program then has its own governance framework and mechanisms designed to carry out that program's goals. For example, "In the OGC Standards Program the Technical Committee and Planning Committee work in a formal consensus process to arrive at approved (or "adopted") OGC® standards,"[362] while the "Interoperability Program is a series of hands-on engineering initiatives to accelerate the development and acceptance of OGC standards."[363]

The basic model for standards development is for a group of OGC members to submit a proposal to start a Standards Working Group for the specific technology and version (e.g., OGC Sensor Observation Service version 2.0). The group forms, subject to review by the OGC Architecture Board and a vote by the OGC members. The SWG does its work, and when they are ready (a few months to a few years later) the completed specification documents are submitted for adoption vote by the OGC members.

An alternate path is also possible, when the standards development effort was done outside OGC. In this case, the specification documents would normally be reformatted to be more consistent with OGC–native standards, and the proposal would be presented as a Request For Comment (RFC) to the OGC members, following the OGC Architecture Board's review. The most notable example of this being done was to get OGC branding for CF/NetCDF, which is developed outside the OGC (Figure 24).

[361] David Arctur, personal email, June 6, 2012.
[362] "Programs," Open Geospatial Consortium, last modified 2012, http://www.opengeospatial.org/ogc/programs.html.
[363] Ibid..

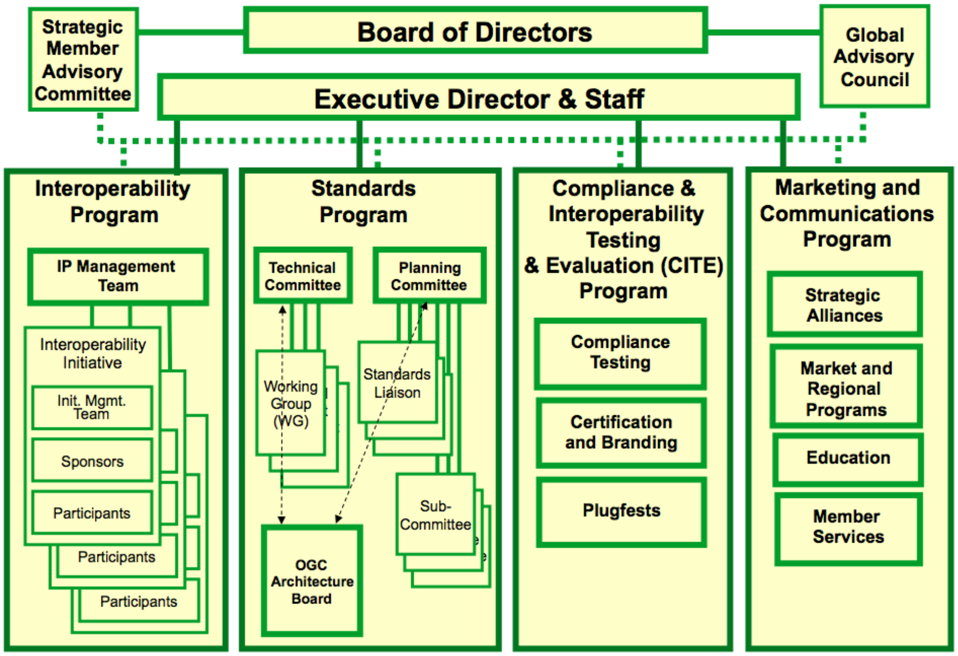

FIGURE 24: OGC ORGANIZATION AND PROGRAMS

OGC has a set of Policies and Procedures to fulfill its mission and vision:[364]

1. Current Bylaws
2. Principles of Conduct
3. Technical Committee Policies and Procedures document for Standards Development.
4. Legal Notices & Term and Conditions
5. Intellectual Property Rights (IPR)
6. Interoperability Program Policies, Processes, and Procedures (IP PP&P)
7. Policies & Procedures for Translation of OGC Public Documents
8. Policies & Procedures for the OGC Architecture Board (OAB)
9. Policies & Procedures for the Compliance Testing Program (CITE)

Staff. "Director, Interoperability Programs" are the program & project managers for Interoperability Program initiatives. OGC staff cannot be working group or subcommittee chairs. At times OGC hires a consultant to lead an interoperability initiative, depending on resourcing and expertise available. These are drawn from a pre-qualified group called the IP Pool (http://www.opengeospatial.org/projects/ippool). Some are contributed by member companies; some are paid by project funds.[365]

LARGE FACILITIES: GENI, IOOS COASTAL AND OCEAN MODELING TESTBED

GENI – GLOBAL ENVIRONMENT FOR NETWORK INNOVATIONS

General Background.[366] GENI is a unique virtual laboratory for at-scale networking experimentation to envision and create new possibilities and future internets. The GENI mission is to open the way for transformative research

[364] "Policies," Open Geospatial Consortium, last modified 2012, http://www.opengeospatial.org/ogc/policies/.
[365] David Arctur, personal email, June 6, 2012.
[366] Hood, Carroll, personal email.

at the frontiers of network science and engineering and to inspire and accelerate the potential for groundbreaking innovations of significant socio-economic impact. GENI, a virtual laboratory for exploring future internets at scale, creates major opportunities to understand, innovate, and transform global networks and their interactions with society. Dynamic and adaptive, GENI opens up new areas of research at the frontiers of network science and engineering and increases the opportunity for significant socio-economic impact by supporting at-scale experimentation on shared, heterogeneous, highly instrumented infrastructure; enables deep programmability throughout the network, promoting innovations in network science, security, technologies, services and applications; and provides collaborative and exploratory environments for academia, industry, and the public.

Funding. GENI is supported by NSF. As of January 2011, awards were made to 83 academic/industrial teams for various projects to build, integrate, and operate early prototypes of the GENI virtual laboratory.

Governance. GENI is a top-down governance model managed by the GENI Project Office (GPO) operated by Raytheon/BBN. The GPO provides system engineering and project management expertise to guide GENI's planning and prototyping efforts. As the management entity for GENI the GPO has responsibilities in several functional areas, as well as responsibilities for development of policies, procedures, and processes that support the ongoing operation of the project. Key among the functional area responsibilities are administration, contracting, financing and budget, legal, systems engineering, project operations, and communications-liaison with many communities of interest to the project.

The overall design of the GPO is such that it is able to effectively communicate and work with a broad range of people and organizations – from government to sponsors of the project, to research and educational users of the GENI facility, to groups federated with the project, to the general public. Note that one such relationship is that between GENI facility developers who are subcontracted to construct the GENI facility. The GPO also provides continuing communications with the computer sciences and networking communities, as represented by the GENI Sciences Committee (GSC) and researchers in universities across the U.S. and abroad.

Further, the organization is guided by a set of policies and procedures that enable it to function on a day-by-day basis, as well as in periods of crisis where rapid decision-making is essential. No part of the management organization is ever allowed to be isolated from the rest, and this means that not only is the organizational structure to be "transparent" to all involved in it but that the inner workings of each functional area are sufficiently visible for collaboration, cooperation, and openness to prevail.

Finally, the organization of the GPO is flexible and readily adaptable to change. The GPO was established during the planning stage of the project, continues through a 5-to-7 year construction stage, and then moves into full operations for a period that could be as long as 15-20 years. During these stages and over this period of time, the GPO will be required to adapt to entirely different environments. It is essential that this adaptability requirement be carefully considered at the earliest stages of GPO design and deployment and built into the structure and processes of the GPO from the start.

GENI encompasses four working groups: Experiment Workflow and Services; Control Framework; Campus Operations, Management, Integration and Security; and Instrumentation and Measurement. Working Group governance is, in general, through rough consensus as judged by the Working Group (WG) chair. WG System Engineer acting as the representative of the GENI Project Office will be responsible for helping the working group understand the needs and constraints of the project office, and the working group is expected to be responsive to these needs and constraints in their activities. In the event the GPO modifies or rejects a WG recommendation, the rationale will be provided to the WG.

WG operations, decisions, and rationales are transparent and are recorded in meeting minutes and open archives. Experience has shown that sometimes a topic may be sufficiently rich or complicated that technical work is best performed by one or more small design teams. These teams may organize how they choose and, in particular, they may not have open participation. However, to be considered for inclusion in GENI, design teams must bring their results to a working group and publish documents through the GENI process. Design team input shall be considered with the same weight as any other working group contribution.

WGs are overseen by the GENI Engineering Architect (GEA). Oversight includes approval of charter creation and changes, staffing decisions, breaking deadlock, defining and/or approving WG operational policies, and ensuring those policies are followed. The GEA is responsible for seeing that a comprehensive and coherent technical design is delivered to the GENI Project Director (GPD) and, thus, will direct working groups. Working group charters (and, hence, working group existence) will be reviewed annually by the GEA and GPD. The GEA and GPD may also perform this review more frequently, although this is not expected to be a common occurrence.

Staff. The GENI Project Office is currently staffed by 15 employees.[367]

SURA US IOOS COASTAL AND OCEAN MODELING TESTBED

General Background. The key elements of the SURA U.S. IOOS Super-Regional Modeling Testbed governance that have been crucial to success are:

1. The appointment of a non-conflicted Testbed Advisory and Evaluation Group (TAEG) to provide objective, independent guidance, and assessment of progress
2. Distributed leadership: each team had a team lead and held regular telecons/meetings of all participants
3. Web site to enable knowledge management and facilitate communication
4. Centralized management and project coordination by SURA

Funding. The SURA U.S. IOOS Super-Regional Modeling Testbed is funded by the National Oceanic and Atmospheric Administration (NOAA).

Governance. SURA's overall management and project coordination for the Testbed not only ensured that the federal funds were expended in a responsible and accountable manner but also includes the full support of SURA's financial and accounting services, grants and contracts management services, and support services. The Testbed enjoyed the full endorsement of the SURA Coastal and Environmental Research Committee (CERC). Although not exercising day-to-day project management, the CERC is the "authorizing" entity for SURA Coastal Research projects. The organization of the Testbed is depicted in figure 2 below.

The Testbed Advisory and Evaluation Group (TAEG) are an independent, technical and scientific advisory group comprised of scientific and computer experts from academia and the federal government. The TAEG makes objective recommendations regarding team selection, resource allocations, and progress assessments and provides a long- range vision and Science Plan for the project.

The Testbed Principal Investigator (PI) works closely with the SURA Project Director who in turn reports to the SURA CEO. The Testbed PI is responsible for leadership and multi-institutional coordination for the Testbed, interacting with the TAEG as necessary to assess project status and develop a long-range vision. The Project Coordinator at SURA is responsible for the day-to-day activities and communications among Testbed PIs and teams and also is the liaison among the SURA HQ, the U.S. IOOS Program Office and the Testbed PIs and teams.

The Testbed is comprised of three modeling teams (Inundation, Estuarine Hypoxia, and Shelf Hypoxia), each with a team leader and a Cyber-infrastructure team which develops the overarching and sustaining Testbed framework.

[367] The GENI Project Office (GPO), http://www.geni.net/?page_id=26, accessed 6/6/12.

Staff. The Coastal and Environmental Research Committee and the Testbed Advisory and Evaluation Group are both volunteer groups. All other participants (with the exception of the IOOS Program Manager) are supported directly or indirectly through external funding (Figure 25).

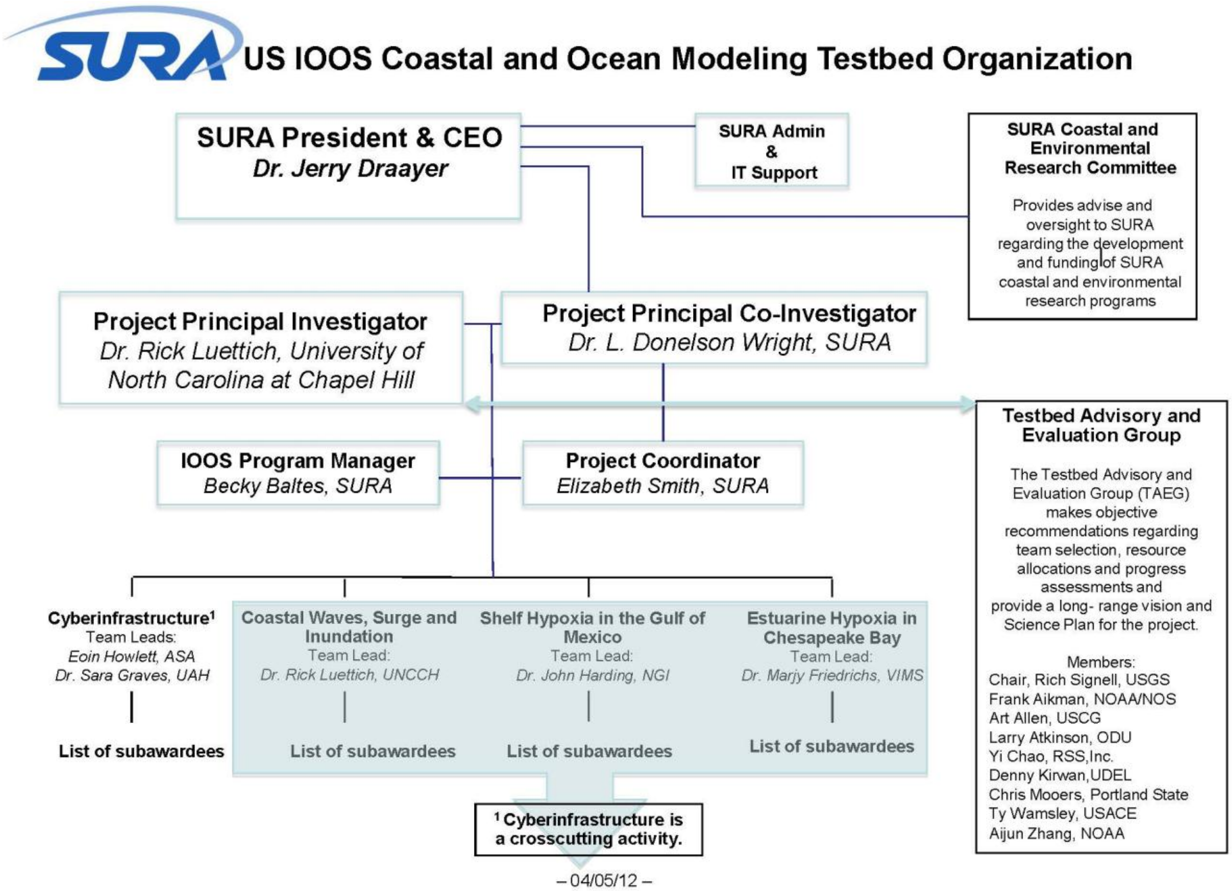

FIGURE 25: SURA US IOOS COASTAL AND OCEAN MODELING TESTBED ORGANIZATION[368]

FUNDING MODELS

The following table summarizes the funding models of various domain science, IT, Federated, and Large Facilities groups, some of which were reviewed previously as case studies. Funding mechanisms of the governance framework will help determine the final structure; however, that structure must remain agile enough to adapt when and if a funding stream is exhausted (Table 19).

Common Funding Models

Category	Program	Governance Funding Model

[368] Courtesy of Gary Crane, personal email, Jun 5, 2012.

Domain Science	Unidata, IRIS, IEDA	Core Grant
	UCAR, CUAHSI	Membership Fee
IT Groups	W3C	Membership Fee
	Apache	Foundation(donations)
Federated Groups	OGC *ESIP Federation*	Membership Fee, Grants for Interoperability projects *Grants*
	DataONE	Core Grant
Large Facilities	OOI, EarthScope, NEON, XSEDE, IOOS	Core Grant

TABLE 19: SUSTAINABILITY/FUNDING MODELS FOR DOMAIN SCIENCE, IT GROUPS, HYBRID GROUPS, AND LARGE FACILITIES.[369]

GEOINFORMATICS AND CYBERINFRASTRUCTURE

EarthCube governance will likely encounter a number of challenges related to IT, scientific vision, leadership, community-building, end-user engagement, and the complexities associated with data sharing and management, among many other challenges. Recent geoinformatics and CI-related workshops and reports have begun to address some of these issues. The wealth of information generated by these workshops will be a valuable resource for EarthCube and are included in this initial governance research review. Geoinformatics and CI-related reports reviewed for the EarthCube governance roadmap include the *Earth and Space Informatics Summit* (2008) and the *Workshop on Working towards a National Geoinformatics Community (NGC)* (2010).

EARTH AND SPACE INFORMATICS SUMMIT

The Earth and Space Informatics Summit (ESI, March 2008) established a basis for better mutual understanding among earth and space science leaders, and confirmed the need for greater collaboration on data issues that require a global outlook. The ESI Summit found that the huge growth in the informatics field requires greater communication and collaboration among globally distributed groups to promote uniformity in practices and standards and to avoid duplication.

Because leadership and governance have a role in how data and information efforts respond to user/stakeholder needs, a balance of leadership and governance is required to successfully advance the geoinformatics agenda. Thus, leadership and governance must be able to cross disciplinary boundaries. Leadership and governance (L & G), or some combination of the two, are vital in different ways, depending on the situation (Table 20).

[369] Mohan Ramamurthy, "Governance Examples" (presentation, EarthCube Governance Virtual Plenary Workshop), April 11 and 17, 2012.

Leadership and Governance

LEVEL	L & G STATUS	Examples
Working level	L, self-G	Marine Metadata Initiative
National/ regional societies	L, what is role for G?	AGU, EGU, AOGS
'Mission' and 'Production' agencies	G, what is role for L?	BGS, USGS, ESA
Regional programs	some L, no G?	EGSO, INSPIRE US National Geoinformatics System GeosciNet Geochemistry Data Network
Global programs	some L, no G?	GEOSS, GMES, GCOS, OneGeology, IPDA
International Association and Union	some L and some G but not uniform	IAGA, IAU, IUGS(CGI), IUGG(UCDI)
International alliances		IVOA, CEOS, Geoscience Info Consortium
Global, inter-union	G, need L	ICSU, WDC, GEO CODATA, WGISS

TABLE 20. LEADERSHIP AND GOVERNANCE RECOMMENDATIONS FROM THE ESI SUMMIT[370]

Currently, leadership and governance are lacking in web services' interoperability, IP rights for data and knowledge, and environmental observatories.

The ESI Summit also identified common needs among geoinformatics practitioners, organized into the 'seven C's:[371]

1. Chart (discover and map what's happening)
2. Communicate
3. Coordinate
4. Contract (use expertise and products of others)
5. Collaborate
6. Converge (reduce complexity and proliferation)
7. Seek Consensus

Specific areas of need were organized into awareness, form, cooperation, agreements, best practice and policy, systems and interoperability models, information commons and open access, and incentives for good data practices. The ESI Summit also established common contributions to address these needs:[372]

1. Raise awareness of informatics and encourage engagement
2. Establish a peer community
3. Two-way communication system between individual domains and international community
4. Share lessons learned
5. Engage with agencies and funding bodies
6. Work toward developing, identifying, and implementing universal and community of practice procedures and standards
7. Identify common issues among important professional bodies
8. Share technical expertise, organizational know-how

[370] Earth and Space Informatics (ESI) Summit, 2008, 5
[371] Ibid., 5
[372] Ibid., 7

9. Provide mutual support for one another's programs

The ESI Summit also produced a number of recommendations for the future. The report recommends that the International Conference on the Unity of Science (ICUS) and other scientific unions have a role in fostering report conclusions mentioned above and that cooperation efforts should leverage existing international frameworks. Specific governance recommendations include producing a formal professional structure and governance centered on ICUS, establishing a forum for cooperation, and lobbying to deal with specific issues, exploring the feasibility of a distinct informatics society to serve the emerging peer community, and exploring the possibility of an international institute of informatics.

The report recommends a two-pronged approach to expanding the geoinformatics community by creating new contacts and working with existing/planned structures. Additionally, more governance structures may be added as necessary, and communication gaps should be dealt with. Other recommendations include international agreements on best practices and standards, WDC certification for relevant groups, a generic and agreed-upon ICSU data policy, shared data policy resources, identification of good ideas that are ready to be shared, and open and timely access to data for scientific research.

WORKSHOP ON WORKING TOWARDS A NATIONAL GEOINFORMATICS COMMUNITY (NGC)

The *Workshop on Working towards a National Geoinformatics Community (NGC)* (2010) identified a number of governance issues that plague the geoinformatics community. The goals of the workshop were to foster geoinformatics education and outreach to the community, explore technical aspects such as semantics and interoperability, and discover what community organization is necessary or desirable to further the goals and needs of the geoinformatics community. Impediments to geoinformatics progress include a lack of technical capabilities, funding, trained people, involvement/buy-in, and opportunities for innovation, in addition to a need for facilitators to bring people together, domain provinciality, and the great diversity of interests in the community. Key geoinformatics problems identified by the NGC are centered on data integration, understanding, trust, and data longevity.[373]

Problem 1: I can't integrate what I can't find. Data application producers must make their products known, and data and application consumers must be able to locate, evaluate, and access resources they need. NGC should foster the development and adoption of standards for a domain-wide distributed catalog that could be used in education and training. Additionally, community organization could be improved with topical sessions at professional meetings and with support planning and interoperability experiments, along with documented results.

Problem 2: I can't use something I don't understand. Data and tools should be easily understandable to diverse audiences. There should be greater efforts across the geoscience community to create more complete documentation and explanation of datasets and analytical tools.

Problem 3: I don't want to use something I don't trust. The solution to this problem requires an understanding of best practices at multiple levels because many different processes and mechanisms contribute to the building of trust. Specifications, procedures, and standards for data and analysis tools and software engineering are key. Users must also be able to access and evaluate the quality and accuracy of data and tools, and communities must support these evaluation methods and be better informed by ongoing research. These trust-building processes

[373] Douglas Walker, Linda Gundersen, Lee Allison, Hassan Babaie, Doug Fils, Steve Richard, Ramon Arrowsmith, Cinzia Cervato, and Steve Whitmeyer, "Workshop on Working towards a National Geoinformatics Community (NGC)," USGS Denver Federal Center, Denver, Colorado, September 23-24, 2010, 1.

and mechanisms require a well-trained Geoinformatics workforce in both IT and domain science. Education and training could be a core NGC activity.

Problem 4: I can't use something that isn't there anymore. The duration of data is very important. Data and analysis tools must be available after a project ends, and metadata and QA/QC procedures must be followed and recorded. NGC could work with international standards and procedures groups to promote data security in the long-term.

Community building and coordination will be necessary to resolve the problems of data integration, sharing, longevity, and trust. The authors recommend a community-developed governance framework and hardware and software to promote technical innovation to advance the geoinformatics community.

CONCLUSION: SECTION 3

The case studies demonstrate that governance is as varied as the institutions it serves, even within the same communities. **Case studies include:**

1. Domain science; CUAHSI, IRIS, Unidata
2. IT: W3C Consortium, DataONE
3. Federations: ESIP and OGC
4. Large Facilities: SURA US IOOS Coastal and Ocean Modeling Testbed

Case studies key points:

1. Each organization has a governance framework developed specifically to meet that organization's needs. For example, even within the federated organizations, OGC's governance framework differs greatly from that of ESIP.
2. Organizational goals vary greatly from organization to organization, as do the management structures and decision-making processes needed to achieve those goals.
3. Funding mechanisms and paid staff vs. volunteers differ greatly from organization to organization.
4. The effectiveness of the governance framework should be evaluated against how well the framework enables the organization to move forward in achieving its goals.

Geoinformatics Reports key points:

Earth and Space Informatics Summit

1. **A balance of leadership and governance** is required to successfully advance the geoinformatics agenda because leadership and governance have a role in how data and information efforts respond to user/stakeholder needs.
2. **Seven geoinformatics governance needs**:
   a. Chart (discover and map what's happening)
   b. Communicate
   c. Coordinate
   d. Contract (use expertise and products of others)
   e. Collaborate
   f. Converge (reduce complexity and proliferation)
   g. Seek Consensus
3. **Common contributions to address these needs:**
   a. Raise awareness of informatics and encourage engagement
   b. Establish a peer community
   c. Two-way communication system between individual domains and international community
   d. Share lessons learned

e. Engage with agencies and funding bodies
f. Work toward developing, identifying, and implementing universal and community of practice procedures and standards
g. Identify common issues among important professional bodies
h. Share technical expertise, organizational know-how
i. Provide mutual support for one another's programs

Workshop on Working towards a National Geoinformatics Community (NGC)

1. **Four major geoinformatics problems:**
    a. I can't integrate what I can't find
    b. I can't use something I don't understand.
    c. I don't want to use something I don't trust.
    d. I can't use something that isn't there anymore.
2. **Community building** will address issues of data integration, sharing, longevity, and trust.
3. **Recommends** a community-developed governance framework.
4. **Recommends** hardware and software to promote technical innovation to advance the geoinformatics community.

Conclusions from Geoinformatics Reports

1. There are actively engaged end-users in the geoinformatics community.
2. Best practices and lessons learned from these reports and others offer a wealth of knowledge from which to draw recommendations for EarthCube governance.

SECTION 4: EARTHCUBE GOVERNANCE RECOMMENDATIONS FROM RESEARCH REVIEW

Presented here are specific governance recommendations from NSF EarthCube white papers from the Governance and Design categories, submitted fall 2011. Both the governance and design category white papers offer a wealth of ideas and recommendations regarding EarthCube governance. Themes from the governance white papers include leadership, purpose, community-building, commodity governance, public-private partnerships, diverse representation in governance framework, open source software, and open science, along with a series of case studies on ESIP, IRIS, OGC, UNAVCO, Unidata, and WInSAR Consortium. Themes from the design white papers include a focus on long-term goals, communities of interest, data trust, knowledge management, space-time location, centralized database, professional networking services, an ECNetwork, scientific forecasts, communities of practice (divided into informaticists, data generators and synthesizers), in addition to case studies on IRIS, ESIP Discovery Cluster, USGIN, USGS, Unidata, QuakeSim, CUAHSI-HS, DataONE and CZO Data. Themes are organized alphabetically by author.

NSF WHITE PAPERS: GOVERNANCE CATEGORY

THEMATIC RECOMMENDATIONS

PURPOSE, LEADERSHIP AND AUDIENCE

Dave Fulker argues in "EarthCube Leadership" that EarthCube must have a "clear, measurable definition of success [that] will help create a framework for accountability and effective leadership."[374] There must be quantitative measures of success (the number of community members positively affected by EarthCube, for example) and qualitative measures ("Incorporat[ing] the ease by which knowledge may be constructed from observed or synthesized data that fall within EarthCube"[375]). In this way, "…beneficiaries of a successful EarthCube would spend less of their time solving problems in data access and use, and they would have fewer misunderstandings about the data they employ." [376]

There must also be a clear assignment of leadership, and leadership ambiguities must be avoided. Fulker argues, "Though I value deep community *engagement* in governance, I think EarthCube will yield diffuse, unimpressive, common-denominator outcomes unless its core leadership responsibilities are clearly assigned to a carefully selected organization."[377] Ideally a single entity would responsible for EarthCube success, and NSF should support the leadership model and funding mechanisms, especially at the beginning of EarthCube. Funding should be "contingent on recipient commitments to cooperate with the lead entity and to comply with EarthCube standards and policies, which presumably will be maintained and promoted by this organization."[378]

The EarthCube audience must be specified and have a high degree of community buy-in. Fulker recommends that the EarthCube audience be focused on academia, including all college/university students, educators, and researchers in all geoscience disciplines and that the audience should be viewed as both data creators and data users.

---

374 Dave Fulker, "EarthCube Leadership," 1.
375 Ibid. 2
376 Ibid.
377 Ibid.
378 Ibid.

COMMUNITY-BUILDING

Bruce Caron argues that community-based governance is the best model for EarthCube. Individual members who have a greater stake in the community as a whole are more likely to contribute to the advancement of the community and be involved in the decision-making process. By reaching out to target populations and actively building a democratic community, the EarthCube community as a whole becomes greater than the sum of its parts. "Community members will be more capable of knowing how to wield collective intelligence effectively to deliver Earth science data and information and to be more motivated to share their resources." [379]

Caron argues that it is crucial for governance structures, including software and democratic practice, to be in place from the very beginning. A number of key questions must be addressed in tandem with setting up the governance framework:

1. How is community created and when should this process begin?
2. What does it mean to be a peer (i.e. if someone is considered a peer, they should have equal access to information and decision-making)?
3. When and how is feedback collected?
4. What processes are in place to respond to criticism and feedback?

COMMODITY GOVERNANCE

Cecelia DeLuca argues in "Commodity Governance" that EarthCube is a community of communities. DeLuca defines communities of communities as entities that have national and international cooperation, distributed resources across institutional and international boundaries, multiple tools and standards in place, multiple agencies and organizational missions, goals and target communities, and no single project manager. Communities of communities require a re-envisioning of governance, promotion, collaboration and communication, and are "characterized by a need for complex, integrated deliverables in a highly distributed development environment, with multiple leads of comparable influence."[380] Two concrete steps can aid in the development of communities of communities:[381]

1. **Use of a mediator, such as a Change Review Board**, that meets frequently and shares the responsibility of decision-making and goal-creation, in addition to prioritizing development tasks, setting contents of releases and schedules for collaboration, and drawing a clear line between the sponsoring institution and how funds are used.
2. **Cyberinfrastructure that organizes contributions and development from different organizations**, and maintains standard governance practices to the extent that they are necessary. Specifically, these tools can "commoditize" and encode governance practices into software development and analysis environments.

PUBLIC-PRIVATE PARTNERSHIP

Kumar et al. in "A Public Private Partnership Opportunity for NSF Earth Cube to Serve Stakeholders and Decision Makers" recommend public-private partnerships for EarthCube. They argue that a public-private partnership, coordinated across government, academia, and industry, is needed to understand both the science user requirements and the stakeholder requirements of EarthCube and that a public-private partnership would effectively serve national, regional, and local decision-makers, the private sector, and the general public. Specific stakeholder roles include:

---

[379] Caron, "Realizing Expectable Returns on EarthCube Investments in Community Building and Democratic Governance," 5.
[380] DeLuca, "Commodity Governance for Communities of Communities," 1.
[381] Ibid., 2.

1. **Government:** facilitate and fund larger projects
2. **Academia:** develop models and methodologies of complex environmental systems and enhance predictive analysis capabilities
3. **Industry:** promote understanding of end-user requirements across economic sectors

EarthCube should produce a flexible and reusable system that can be extended by future organizations which were not necessarily involved in the original design process. Goals of an EarthCube public-private partnership including the following:

1. Enhanced access and expanded scope of support to end users
2. Tailored process for delivery of information to meet users' needs and practices
3. Bridging the gap between scientists and decision-making organizations through standards for sharing data
4. Determination of what new data are needed to meet end-user needs
5. Effectively serving national, regional, and local decision-makers, the private sector, and the general public

DIVERSE REPRESENTATION IN GOVERNANCE FRAMEWORK

Idaszak et al. in "Insights and Precepts from a Workshop on Creating a Software Institute for Environmental Observatories" argue for diverse representation of all stakeholders in the EarthCube governance framework. They argue that an environmental observatories institute could implement a governance structure that is representative of the community across multiple dimensions including observatory type, codes, nature of research processes, education, analytics, and information technology. There should be diverse, multi-sector participation with comprehensive expertise in the represented areas. Governance should be driven by, and be responsive to, the community that the software is intended to serve.

The authors envision an effective, distributed network infrastructure that is shared, interoperable, extensive, and modular that fosters best practices in policy, governance, and virtual organization. Specifically, representatives from all environmental observatories would be included in the governance framework, and technical personnel would serve on the development team. End-users would determine the direction and requirements that the software ecosystem would address.

OPEN SOURCE SOFTWARE GOVERNANCE AND OPEN SCIENCE

Bedard et al. in "What Open Source Software Development Teaches Us about Scientific Governance" argue for open source software and open science as a model for EarthCube. They argue that existing governance models for scientific research are not equipped to address the Grand Challenges (climate change prediction and mitigation, managing greenhouse gases, and hazard analysis and management) and that the Grand Challenges will require "transformative discovery and innovation across disciplines; advanced computational models and algorithms that operate at multiple scales; and the collection, management, and analysis of massive, heterogeneous datasets."[382] Specific shortcomings of existing systems include static collaboration, review lag, misplaced incentives, and unsustainability.

Open source software addresses these shortcomings through and contributes knowledge to the commons while preserving constructive competition. Policies to facilitate open science in EarthCube may include funding to permit "rapid reconfiguration" of project teams throughout the research process as new priorities emerge and the development of data and software in anticipation of direct peer review and collaboration. The authors also

[382] National Science Foundation, 2011, quoted in Bedard, et al., "What Open Source Software Development Teaches Us about Scientific Governance," 1.

recommend a discussion of open problems, which is another important aspect of open governance. These problems include determining the right funding mechanisms, how to avoid "infinite loops," determining when research should be closed, making education and outreach a priority, and determining the use and implementation of metrics.

Finally, open source governance has many short-term and long-term advantages. In the short-term, open source governance leads to self-sustaining software projects, thus producing "continual returns to the originating community as they evolve in concert."[383] In the long-term, they will "change the culture of the community and the NSF to embrace sustainable open source and software engineering practices," resulting in "unforeseen innovation" and leading to "new open science that is continual and sustainable."[384]

TOPICAL RECOMMENDATIONS/CASE STUDIES

FEDERATION OF EARTH SCIENCE INFORMATION PARTNERS (ESIP)

Chris Lenhardt in "Governance or Process: Building Community through Collaboration for Shared Solutions" presents the Federation of Earth Science Information Partners (ESIP) as an existing organization whose community-based governance model fosters collaboration and coordination of interoperability efforts across the geosciences community. ESIP's governance model is discipline-neutral; it allows IT experts and scientists to work to improve discovery, access, integration and interoperability, it connects distributed and heterogeneous communities, and it employs broad community technical expertise. ESIP optimizes collaboration through in-person and virtual meetings, which enhance participating members' capabilities, and has in place neutral forums for knowledge exchange and collaboration.

INCORPORATED RESEARCH INSTITUTIONS FOR SEISMOLOGY (IRIS)

Stump and Simpson present the Incorporated Research Institutions for Seismology (IRIS) Consortium as a model for EarthCube. IRIS is directed by a community with a wide range of research interests. The governance framework establishes community ownership of instrumentation resources and data, and the framework is designed to allow IRIS to adapt to changes in technology and in the growth of the Consortium. In addition, IRIS functions as an interface between the scientific community, funding agencies and the IRIS programs. IRIS programs and facilities are multi-user and multi-use.

IRIS is governed by Consortium members (114 U.S. research universities and research institutions). Standing Committees (USArray Advisory, Planning, Program Coordination, Instrumentation) and Board of Directors and Subcommittees make decisions. In addition, IRIS employs workshops, annual meetings, symposia, and newsletters to promote on-going interactions with scientists and member institutions and to receive feedback. These mechanisms allow the research community to communicate its evolving needs to funding agencies.

OPEN GEOSPATIAL CONSORTIUM (OGC)

Arctur et al. in "Governance: An OGC white paper for NSF EarthCube" recommend OGC governance mechanisms as a model for EarthCube as "a proven framework of tools, policies, and procedures for governance of long-term, interdisciplinary and multi-sector collaboration for high-impact scientific research for societal benefit."[385] Arctur et al. argue that good governance must be able to:

1. Describe complex systems
2. Coordinate across institutional boundaries.

[383] Bedard, et al., "What Open Source Software Development Teaches Us about Scientific Governance," 7.
[384] Ibid.
[385] Arctur et al., "Governance: An OGC white paper for NSF EarthCube," 1.

3. Create working groups
4. Mobilize research findings and apply to broader community
5. Create compliance guidelines
6. Advance lifecycle of best practices and standards
7. Consider culture, society, law, trust, uncertainty, and "authentication and authorization for discovery and access"[386]

OGC has a number of mechanisms in place that promote good governance, including:

1. Public and private virtual teams (public websites, members-only portals, members-only wiki spaces, and a public wiki space)
2. Policies and Procedures (P & P) in new standards adoption that reduce competitive pressures among participants, protect intellectual property rights, and protect privacy
3. Interoperability program with separate P & P
4. A strong network of IT standards organizations

The authors propose to focus on outreach and education to increase OGC's capacity to support more scientific domains and open science.

UNAVCO

Holt and Miller in "*Community governance: UNAVCO Facility models for EarthCube implementation"* present UNAVCO as a successful example of "how flexible governance structures can be optimized to project needs and science requirements using science community oversight."[387] UNAVCO is an academic consortium composed of 98 universities and 71 international and other organizations. Its purpose is "….to support and promote scholarly research involving the application of high precision geodetic and strain techniques such as the Global Positioning System (GPS) to Earth science and other related fields of study."[388]

Holt and Miller argue, "The key to effectiveness of a community governance structure is clear articulation of community mission, vision, and strategic goals."[389] UNAVCO serves as a mechanism to develop programs and facilities to support the science vision of its community, in addition to functioning as an interface between the scientific community, funding agencies, and UNAVCO programs. Specifically, "facility managers respond to community needs using methods and frequency tailored to specific programs or projects and to specific science requirements within the community."[390]

UNAVCO has created an effective community governance framework through a formal process of strategic planning, involving the Board of Directors and community stakeholders, in addition to various Standing and Advisory Committees. Community involvement supports broad participation and effective oversight of UNAVCO programs.

UNIDATA

Mohan Ramamurthy in "Unidata Governance: A Quarter Century of Experience" presents the Unidata governance framework as a model for EarthCube. Effective governance of EarthCube should actively engage diverse users; provide strong leadership and oversight of the project to foster close collaboration, coordination and development

[386] Ibid.
[387] Holt and Miller, "Community Governance," 4.
[388] Ibid., 2.
[389] Ibid., 1.
[390] Ibid., 3.

among distributed development activities and EarthCube groups; facilitate alignment of program plans and priorities to meet community needs; and foster the successful execution of the EarthCube mission.

Unidata's governance framework consists of a "centralization only of selected functions and overall leadership, a high degree of community engagement, and strategic partnerships between agencies and universities."[391] Unidata is governed by a Policy Committee and a User's Committee, along with other interim groups formed on an as-needed basis. The Policy Committee advises on the direction and program priorities, in addition to making sure university community needs are met. The Users Committee seeks user input and feedback.
Unidata's governance framework focuses on assessing and responding to community needs. This includes advocating and negotiating agreements on behalf of the community on data issues, fostering community interaction to promote sharing of data, tools, and ideas, and consensus building.

Unidata employs a number of governance mechanisms to bring its governance framework into practice. These mechanisms include active, independent, and informed oversight and stewardship; consensus-oriented and pragmatic approaches to decision-making; balanced representation from the full range of participating institutions; shared responsibility between the governing committees and the program with a clear delineation of the roles; direct lines of communication between governing committees and staff: and focus on long-term strategic direction with nimbleness to respond to rapid changes.

WINSAR CONSORTIUM

Fielding et al. in "Challenges Facing the WInSAR Consortium from the Coming Tsunami of SAR Observations" present the Western North America InSAR Consortium (WInSAR) governance model as a low-cost and effective way to achieve both EarthCube and WInSAR goals. WInSAR serves as the primary repository of Inferometric Synthetic Aperture Radar (InSAR) data for the solid earth community in the United States. As of fall 2011, it had 96 member institutions elected and administered by UNAVCO, which provides administrative and logistical support to the Consortium.

An executive committee, whose members are elected bi-annually by Consortium members, ensures operations meet policies set by WInSAR membership and sets operational goals, such as data purchases, distribution, archiving, and data products. WInSAR holds teleconferences, whose notes are posted on the UNAVCO website, and holds an annual business meeting to learn about resources and discuss strategies. Community input from emails and other communication is collected throughout the year. Annually, WInSAR holds an InSAR for Beginners Short Course for interested WInSAR members.

EARTHCUBE NSF WHITE PAPERS: DESIGN CATEGORY

The NSF Design White Papers offer a number of useful insights on governance framework, overall design, and sustainability. Design White Papers are organized thematically, and although some Design White Papers focus on organizations previously mentioned in Governance White Papers, they offer distinct, design-related information, thus are kept in the design section. Design White Papers that did not specifically include governance components are left of out this review.

THEMATIC RECOMMENDATIONS

FOCUS ON LONG-TERM GOALS

[391] Ramamurthy, "Unidata Governance," 6.

Ahern and Koper in "Flexibility in an EarthCube Design Approach: Ideas from IRIS and its Partners" argue that governance must advance the interests of its intended users and must suit long-term EarthCube goals. EarthCube must evolve to meet changing needs and opportunities, and form should follow function: Forms of governance and management should be tightly linked to functions that implement EarthCube goals. Ahern and Koper suggest that academia should provide leadership, while potential governance models for EarthCube include NSF Directorate (guided by a FACA Advisory Committee), Consortia (such as IRIS, UNAVCO or UCAR), or a community-organized model described as a "non-corporate amalgamation of separately funded projects,"[392] like GeoPrisms.

Ahern and Koper argue that sustainability of EarthCube is linked to the intrinsic value of investments in data collection and preservation. Thus, EarthCube must "demonstrate how sustained investments in data resources contribute to advances in knowledge and the public good."[393] Additionally, Federal agencies (such as NSF) are responsible for long-term data collection and preservation, although public-private partnerships also have potential in fostering EarthCube sustainability.

COMMUNITY-BASED GOVERNANCE

Bowring et al. in "Building a Software Engineering and Informatics Community of Practice and a Software Foundation for the Advancement of EarthCube" recommend a community-based discovery governance model, in which "eventually community-supported standards will emerge via collaborations."[394] They point to the W3C Consortium as a good example. The development processes should be transparent, responsive to community input, serve as teaching tools, and should be based on community feedback. They suggest that ongoing consideration and study sustainability should be a requirement for future funding.

Braeckel et al. in "EarthCube Design Approach" argue that EarthCube will be formed by Communities of Interest (COI), defined as groups of users who share a common interest within EarthCube. Governance must promote the collaboration of users, processes, and resources inside and outside of the geosciences. Users must also be able to trust in the utility, quality, and adaptability of the system. Predicted governance challenges include fitting communities that internally manage their resources into the larger EarthCube community, taking advantage of existing investments and capabilities to encourage collaboration and commonality, and integrating existing resources and services in an inclusive manner. Braeckel et al. illustrate their EarthCube vision in Figure 26 below.

---

[392] Ahern and Koper, "Flexibility in an EarthCube Design Approach" 3.
[393] Ibid., 10.
[394] Bowring, Bowring, and Walker, "An Earth Cube Design Approach," 2.

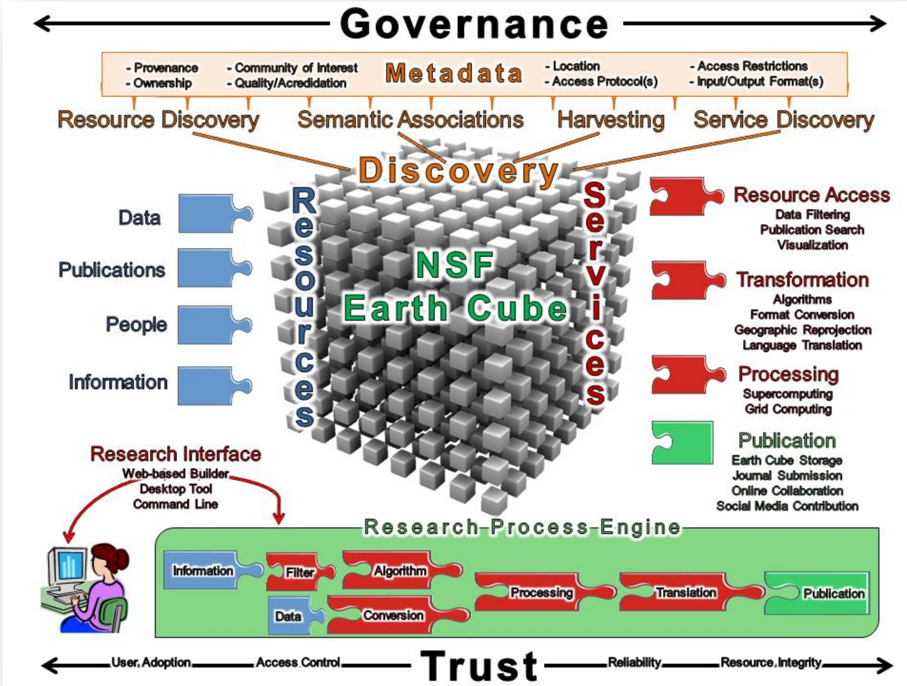

FIGURE 26: EARTHCUBE VISION[395]

If EarthCube is to be sustainable, governance framework must anticipate and allow for the evolving needs of the EarthCube community. Braeckel et al. argue that the user community itself is the most valuable source of innovation, yet "EarthCube must remain stable while supporting change."[396] Lastly, for EarthCube to be sustainable, there must be trust in the system as a whole.

COMMUNITIES OF PRACTICE: INFORMATICISTS, DATA GENERATORS, SYNTHESIZERS

Peters et al. in 'Towards a New Distributed Platform for Integrative Geoscience: An EarthCube Design Approach and Prototype Plan" recommend a distributed, internet-like system for EarthCube, built on lessons learned, "interfering as little as possible with the geoscience communities that have already made important and successful decisions about how to manage and distribute their data."[397] The authors divide the geosciences community into three main communities of practice. These communities include Informaticists (scientists and computer scientists with IT and knowledge management expertise), Data Generators (researchers that collect and use observational and experimental data), and Synthesizers (community of scientists exploring research questions requiring data synthesis and analysis).

Peters et al. recommend first engaging the Synthesizers for several reasons: they are already pre-adapted and motivated to learn good practice: they form a heterogeneous community in terms of skills, tools and backgrounds: they provide a good opportunity to improve knowledge production and transmission: they can be intermediaries between Generators and Informaticists; and many Synthesizers are early to mid-career scientists familiar with social networking and other technologies. Specific steps to engage the Synthesizers' community of practice include:

---

[395] Braeckel et al., "EarthCube Design Approach," 4.
[396] Aaron Braeckel et al., "EarthCube Design Approach."
[397] Peters et al., 'Towards a New Distributed Platform for Integrative Geoscience," 5.

1. Strengthen existing communities of practice (improving connectivity among and within communities)
2. Use local knowledge to raise awareness of good practices and success stories (solutions from one area could work well in others)
3. Strengthen networks to scale up
4. Assume success depends on on-going assessment and adaptation

The EarthCube governance framework should develop basic principles and guidelines for publishing geoscience data and data products to EarthCube. The authors suggest a number of governance mechanisms to foster EarthCube sustainability. These include a "Center for EarthCube Support" to organize operation strategies, short courses and workshops, social networking sites linking people and data, an Informatics Help Desk, and a Pilot Steering Committee.

DATA TRUST

Couch and Bedig in "Building Plausible Representations of Physical Entities" argue that EarthCube design and governance should focus on fostering trust of data. Governance framework must create a broad and shared chain of trust from data providers to scientists to end users of scientific results. Important trust issues that governance must address include: What are trustable data? How do trust and distrust arise? What are the foundations of trust? How can trust be fostered and enabled? Governance can create a shared notion of data plausibility, decide what data sources are considered high quality, and select commonly accepted models to be used for cross-validation.

The authors use a digital watershed as a model to highlight these issues. They argue, "Creating a shared concept of a digital watershed is not a simple matter of combining data, but rather, requires developing a shared scientific concept of data plausibility that does not currently exist," and that it "requires scientific mechanisms for data integration… and…scientific governance of data collection and integration mechanisms to assure plausibility and community usefulness of the results."[398] Couch and Bedig argue that resolving these complex issues will require a combination of governance, modeling, and infrastructure initiatives.

TOPICAL RECOMMENDATIONS/CASE STUDIES

IRIS

Ahern, Simpson and Stump in "Guiding Principles and Best Practices from the IRIS Data Management System" argue that the IRIS Unifying structure encourages operational efficiencies and coordination across the geosciences and is a good model for EarthCube governance. IRIS is composed of one community governed facility that links observational systems, data collection, archiving and distribution. IRIS engages scientific leadership, avoids conflicts of interest, promotes a free and open data exchange, and encourages a culture of data sharing, in which providing data is part of reciprocal exchanges to benefit the broader community.

USGIN

Allison et al. in "Geological Survey Contributions to Building Cyberinfrastructure in the Geoscience" argue for EarthCube to take advantage of the existing U.S. Geoscience Information Network (USGIN) and its tools and applications to understand data. USGIN's goal is to build catalog and web services to link and allow access to holdings of geological surveys in the US, Governance recommendations include expanding the geoinformatics community of practice to establish a long-term governance model to provide a sense of leadership and a source of authority in decision-making and priority-setting. The authors recommend that participants learn, develop, evolve and coordinate network-building together and to increase engagement and participation of state geological

[398] Alva Couch and Alex Bedig, "Building Plausible Representations of Physical Entities."

surveys and communication among the states, USGS, and other stakeholders. Engaging in existing face-to-face meetings such as CID, ESIP, OGC and USGS can encourage adoption of these recommendations.

Allison et al. argue that State geological surveys are government entities that would like to continue their core missions and therefore provide continual data and development for USGIN. Additional groups and companies will also create greater demand for USGIN and will help spread maintenance costs across the geoinformatics community.

USGS

Bristol et al. in "Role of Government Data Agencies, Virtual Observatories, and Communities of Practice in Sustaining EarthCube" recommend that collaboration among government entities and legislative mandates form bases of interaction between government data agencies and EarthCube participants. The authors point to the previous work of the USGS in building and sustaining cyberarchitecture to support its earth science mission and to the National Geological and Geophysical Data Preservation Program (NGGDPP), an existing collaboration between the Department of the Interior and State Geological Surveys. Collaboration among government entities and EarthCube users to create a shared CI vision for the next 10 years is necessary for EarthCube to be effective, efficient, and sustainable. In this way, the EarthCube "community will leverage its resources, find solutions and establish a cyberinfrastructure quicker and more efficiently and possibly move science and innovation forward at a faster pace."[399] Deliberate facilitation of communities of practice to foster collaboration and manage competition toward shared goals is a means to facilitate greater advancements and innovations in scientific research.

UNIDATA

Davis et al. in "An EarthCube Design Process: Unidata's Perspective" base their EarthCube governance recommendations on Unidata's scalable, loosely federated model, and they offer a number of governance-related suggestions to promote EarthCube sustainability. They suggest a framework that promotes strong community buy-in that is built on existing organizational infrastructures and mechanisms for international, public and private collaborations that work to obtain funding from a variety of different sources. Additionally, National and Domain-Specific centers should support, build, and advocate for their communities.

QUAKESIM

Fox et al. in "EarthCube and EarthQuake Science" argue that the Earthquake science field can be a potential partner to help provide the CI envisioned by EarthCube. They point to QuakeSim (a computing infrastructure for fault modeling), whose goal is to help formerly isolated groups work more effectively together. QuakeSim promotes open processes, with a focus on the creation of data products. Goals include greater integration, reliance, and accountability of stakeholders to each other. Solutions are both technical and sociological.

ESIP DISCOVERY CLUSTER

The Earth Science Information Partners Discovery Cluster (ESIP-DC) in "A Lightweight Approach to Earth Science Data Discovery" proposes a lightweight approach to governance, standards, and IT, modeled after the ESIP cluster model. ESIP is composed of a variety of academic institutions, commercial interests, government agencies, and NGOS, all of whom are involved in the earth sciences. It has a well-defined governance structure regarding business activities, and technical progress is made through ESIP "clusters," defined as self-organizing groups that focus on a particular issue. The main goal of the cluster groups is to integrate their findings across ESIP.

399 Bristol, et al., "Role of Government Data Agencies, Virtual Observatories, and Communities of Practice in Sustaining EarthCube."

The ESIP Discovery Cluster is a volunteer-based organization that is consensus driven. The ESIP Discovery Cluster depends on open source software and collaborates to solve problems of interoperability, distributed services, and adoption of different open standards. The Discovery Cluster follows a five-step governance process, in which all steps are posted to a mailing list and/or wiki:

1. Submission of new proposals
2. Forum to review proposals
3. Author revision based on feedback
4. Voting on change proposals
5. Ratification or rejection by editors

The authors argue that EarthCube governance "should allow for emergent governance structures that can leverage the efforts of both volunteer and funded participants."[400] There is also a need for EarthCube governance framework to develop and follow an agreed-upon process to coordinate viewpoints, decisions, and interoperability.

AGORA-NET

Hoffman, Curry and Pu in "EarthAgora: A Collaborative Knowledge Management Tool" present AGORA-net as a model for EarthCube, in which knowledge is graphically organized based on argument maps that link arguments to one another and provide forums for debate, comment, and counterargument. The authors argue that the governance framework should be developed by the EarthCube community itself and that governance must be open, transparent, and inclusive. In addition, EarthCube governance must be able to acquire current and future user input and requirements, respond to changing data and science needs, adapt to and adopt new technologies, coordinate components and facilities, and foster partnerships and community participation. Specific governance mechanisms include an EarthAgora advisory board to design disseminating mechanisms and training opportunities for the community and a "Center for Knowledge Transformation" to serve as the organizational and administrative core for long-term development.

SPACE-TIME LOCATION

Kelbert and Pavlis in "Earth Sciences: The Tower of Babel" argue that a space-time location can serve as common ground for communication among earth scientists. Thus, governance framework must focus on challenges in data integration. The EarthCube CI framework would foster sharing and interdisciplinary collaboration, in addition to maintaining geophysical science information for the future.

CENTRALIZED DATABASE

Leetaru et al. in "EarthCube Design Approach" propose an "open access archive of all publicly available geosciences data in the United States and eventually the world,"[401] whose long-term vision would be led by scientists. Governance mechanisms would include an Executive Committee and a Steering Committee. The authors also suggest a "central master framework to allow for a quantitative decision-making process,"[402] which would provide criteria to holistically evaluate every component of the system, manage risk, ensure integration efforts meet project objectives and schedule, create a strong consensus on data-sharing, and improve the workflow for creating metadata. They argue that a strong governance structure is crucial to EarthCube sustainability, education and outreach, and a variety of potential private and public funding mechanisms.

[400] ESIP-DC, "A Lightweight Approach to Earth Science Data Discovery," 5.
[401] Leetaru et al., "EarthCube Design Approach," 1.
[402] Ibid.

PROFESSIONAL NETWORKING SERVICES

Long and Remick in "A Data-Based Professional Networking System" recommend professional networking services to link scientists to data, with a focus on professional relationships, protocols, formats, services and analytical tools. This system would address current challenges in scientific CI, including an unprecedented generation of data, a need for more interdisciplinary studies, and a need for researchers to be aware of what others are doing to avoid duplication. Governance should be community-based. Each organization would find different ways to organize data and metadata, thus it would be necessary to establish communication to foster professional networking and a semantic roadmap. Flexibility must be built into governance through discussion threads, periodic meetings, and adoption of the most effective standards for researchers who use the system. Sustainability can be achieved through the implementation of best practices for a professional networking service, a semantic mapping service, and an open source model with a governing foundation responsible for oversight.

EC-NETWORK

Mader et al. in "An Architectural Approach for Sharing, Discovery and Knowledge Dissemination on EarthCube" recommend an ECNetwork of interconnected nodes, much like the internet, to build connections among data, resources, and disciplines. The governance framework should support distributed administration and available resources, create flexibility, promote scientific creativity, and take advantage of local expertise. It should nurture open participation, including making educational and other resources easily available and making infrastructure components that are easy to use. Specific governance mechanisms include an overall governing committee or board and working groups/subcommittees to focus on different projects. Potential areas of focus include defining the governance model, defining data and other standards, and implementing key project technologies. Board and working group members should be from the EarthCube community, NSF, and other interested government agencies. The W3C Consortium and Apache Foundation offer potential governance models for EarthCube.

Mader et al. argue that in order for EarthCube to be sustainable, the EarthCube community must have sufficient incentives to bring together resources to improve and maintain the EarthCube system. A non-profit foundation could be established to raise funds to maintain and improve aspects of the system, similar to the role of the Wikimedia Foundation for Wikipedia. In addition, EarthCube must bring value to private industry to help secure funding.

SCIENTIFIC FORECASTS

Philip Maechling argues in "Focus on Forecasts" that EarthCube should focus on scientific forecasts. Specifically, a focus on the organizational and scientific process to help researchers improve geoscientific forecasts would drive EarthCube scientific, organization, and CI development. This would be a bottom-up process, organized around a well-defined purpose, to develop the simplest and most basic capabilities needed by the largest number of groups. Governance would be divided into three groups:

1. **Economic**: Identify economically valuable and scientifically improvable forecasts
2. **Scientific:** Develop forecast evaluation procedures and define processes
3. **Technical:** Oversee computational and data management infrastructure and operations

EarthCube must be scientifically and economically valuable to researchers in order to be sustainable.

DATAONE

Michener et al. in "DataONE-Enabling Cyberinfrastructure for the Biological, Environmental and Earth Sciences" present DataONE as an existing platform that allows scientists to focus on large-scale, long-term problems. DataONE supports rapid discovery of data and world access to data centers, is built on existing data centers, is structured as a global, federated network focused on systems interoperability, and facilitates evolving communities of practices. DataONE employs stakeholder assessment, user scenarios, usability testing, and engagement with domain scientists in use cases to evaluate CI projects. Participatory, user-centered processes are employed for on-going evaluation of DataONE architecture. Specific goals of evaluation processes include:

1. Identification and prioritization of stakeholder communities
2. Development of understanding of their perceptions, attitudes, and user requirements
3. Analysis and assessment of usability
4. Engagement of science teams in grand challenge science exemplars

CUAHSI-HIS

Tarboton et al. in "Advancing Solutions for an EarthCube Design: What Can Be Learned from the CUAHSI HIS experience?" argue that CUAHSI-HIS goals align well with EarthCube. These goals include data access (improving access to a large amount of high-quality data), hydrologic observatories (storing and synthesizing hydrologic data for a region), hydrologic science (creating a stronger hydrologic information infrastructure), and hydrologic education (bringing more hydrologic data into the classroom). Promoting data sharing and interoperability can help CUAHSI achieve these goals and would serve as a means to enable hydrologic analyses that integrate data from multiple sources.

Governance should focus on community process and open development. Specific mechanisms include establishing and documenting a governance model for development activities, creating tools to support community functions and interactions, and active engagement of partners and contributors. This will allow potential contributors to know how data will be handled. The authors note that the volunteer approach of common open source development projects may not be appropriate for the complexity and size of EarthCube. Instead, it will likely be necessary to work with a consortium with a pre-established infrastructure and sense of trust, such as CUAHSI. CUAHSI already has mutually beneficial industrial commercial, agency, and organizational partnerships. The authors recommend that "EarthCube will need to foster these sort of synergies [commercial, agency and organization partnerships mentioned above] and be a software ecosystem to which many can contribute."

CZO DATA

Zaslavksy et al. in "Data Infrastructure for the Critical Zone Observatories (CZOData): an EarthCube Design Prototype" argue that the CZO Data focus on efficient data and knowledge integration to understand and model earth processes is relevant to EarthCube. EarthCube governance should reflect the long-term vision and scope of EarthCube and make clear how components should be managed, how EarthCube operates with infrastructures, and how EarthCube will evolve. There is a potential for overlapping and evolving governance models to be put in place to deal with different aspects of EarthCube governance. The authors also recommend policies for resolving potential conflicts:

1. PI or group of PIs should define governance framework for observatories and individual research sites.
2. Domain CI projects require more complex governance framework. For example, CUAHSI HIS includes a consortium of universities with a governing board, development teams and user committee, an operational and curation support metadata catalog, a community consensus process for domain models, vocabularies, data access protocols/services, and catalogs.
3. Super-domain (or cross domain) infrastructure shares many of the same governance components as CI projects that define curation and management policies, including vocabulary cross-walks, catalog federation, and standard service interfaces.

OGC has a successful governance model in which communities of practice develop and present standards for data and exchange. Working groups define the scope and use cases, how they relate to existing and proposed standards, and how they can be extended. Proposed specifications are shared with larger OGC membership and pass through a series of approvals before the final vote. In this way, an open, transparent and formal process is maintained.

ADDITIONAL EARTHCUBE RECOMMENDATIONS (NON-WHITE PAPER SOURCES)

EARTHCUBE AS A DOUBLE LOOP ORGANIZATION

Bruce Caron argues for EarthCube to be structured as a double loop organization,[403] in which the processes and mechanisms developed to fix certain problems simultaneously solve the larger issues that first created these problems. The governance framework for Double Loop Organizations allow stakeholders/community members to govern their own governance framework by agreeing to rules that solve common concerns related to interactions, such as data sharing (Figure 27).

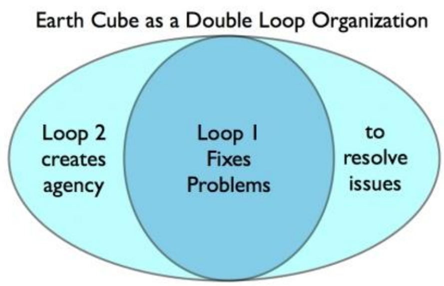

FIGURE 27: EARTHCUBE AS A DOUBLE LOOP ORGANIZATION[404]

The first loop fixes problems of when planning should happen and how. The process begins with discussions about problems to be fixed and their respective solutions, which clarifies the level of governance activities needed to reach these solutions. The first loop then guides the governance framework of the second loop by determining what requirements are needed. For example, if standards are required, the second loop engages to build a strong community. If it is determined that better communication is needed, a weaker community will suffice.

The second loop governance framework is decided by the community and requires much community-wide discussion. The contents of the second loop would include community building, protecting expressive capabilities, voting, officers, working consensus, constitution, bylaws, vision statement, and goals. The second loop creates "agency," allowing community to be owners. Community members have the capability to govern how the community will fix its problems, and are able to answer questions about governance mechanisms and management.

Caron presents ESIP as a good example of a Double Loop Organization. ESIP spent two years to finalize its constitution and bylaws (second-loop outcomes), and the final vote was ratified by consensus. Working and thematic groups (clusters) are also created to fix problems ESIP Members face (first loop).

For this model to work in EarthCube, it is necessary to examine both loops. Loop one could be employed to fix common problems (data sharing and use, for example), and loop two could describe how the EarthCube community can become owners of the governance process. Specifically, loop two would allow EarthCube

[403] Bruce Caron, "EarthCube as a Double Loop Organization," Virtual Democracy, last modified April 4, 2012, http://cybersocialstructure.org/2012/04/04/earth-cube-as-a-double-loop-organization/
[404] Ibid.

community members to determine which fixes for Loop One problems are most important, and community members would be able to determine how to engage the wider community to implement these solutions.

CONCLUSION: SECTION 4

EarthCube White Papers (submitted fall 2011, Governance category and Designs Category) contain several recommendations for EarthCube governance.

Thematic Recommendations

1. **Clarity.** EarthCube governance should make clear how components should be managed, how EarthCube operates with infrastructures, and how EarthCube will evolve.
2. **Collaboration.** Governance must promote the collaboration of users, processes, and resources inside and outside of the geosciences.
3. **Commodity Governance and Communities of Communities**. EarthCube will be a community of communities (entities that have national and international cooperation, distributed resources, multiple tools, standards, agencies, organizational missions, goals and target communities, and no single project manager). Communities of communities require a re-envisioning of governance, promotion, collaboration and communication.
4. **Community.** The EarthCube audience must be specified and have a high degree of community buy-in. EarthCube governance must advance the interests of its intended users.
5. **Communities of Interest.** EarthCube will be formed by Communities of Interest (COI - groups of users who share a common interest within EarthCube).
6. **Communities of Practice: Informaticists, Data Generators, Synthesizers.** The geosciences community can be divided into three main communities of practice: Informaticists (scientists and computer scientists with IT and knowledge management expertise), Data Generators (researchers that collect and use observational and experimental data), and Synthesizers (community of scientists exploring research questions requiring data synthesis and analysis).
7. **Community-based governance.** A community-based governance model is the best model for EarthCube because individual members who have a greater stake in the community as a whole are more likely to contribute to the advancement of the community and be involved in the decision-making process. **Key questions include:**
    a. How is community created and when should this process begin?
    b. What does it mean to be a peer (i.e. if someone is considered a peer, they should have equal access to information and decision-making)?
    c. When and how is feedback collected?
    d. What processes are in place to respond to criticism and feedback?
8. **Diverse Representation in Governance Framework.** There should be diverse, multi-sector participation with comprehensive expertise in the represented areas. Governance should be driven by and be responsive to the community needs and wants.
9. **Evolution.** EarthCube must evolve to meet changing needs and opportunities. The user community itself is the most valuable source of innovation.
10. **Form Should Follow Function.** Forms of governance and management should be tightly linked to functions that implement EarthCube goals.
11. **Goals.** EarthCube governance must reflect the long-term vision and scope of EarthCube, and must suit long-term goals.
12. **Leadership.** There must be a clear assignment of leadership.

13. **Open Source Software Governance and Open Science.** Open source software and open science are models for EarthCube governance. Open source governance contributes knowledge to the commons while preserving constructive competition and will enable scientists to more effectively address the Grand Challenges.
14. **Public-Private Partnership.** A public-private partnership (government, academia, and industry) would effectively serve national, regional, and local decision-makers, the private sector, and the general public.
15. **Success.** There must be a clear definition of success, with qualitative and quantitative measures to evaluate progress towards overall goals.
16. **Sustainability.** Sustainability of EarthCube is linked to the intrinsic value of investments in data collection and preservation. Thus, EarthCube must "demonstrate how sustained investments in data resources contribute to advances in knowledge and the public good."
    - EarthCube sustainability is also dependent on a governance framework that anticipates and allows for the evolving needs of the EarthCube community.
    - There must be trust in the system as a whole for EarthCube to be sustainable.
    - The EarthCube community must have sufficient incentives to bring together resources to improve and maintain the EarthCube system.
    - Sustainability can be promoted through use-cases that allow scientists to imagine how their research can be enhanced through CI.
    - Education and outreach to scientists and the broader EarthCube community can also promote sustainability.
    - Governance mechanisms should foster EarthCube sustainability. A potential mechanism is a "Center for EarthCube Support" to organize operation strategies, short courses and workshops, social networking sites linking people and data, an Informatics Help Desk, and a Pilot Steering Committee.
17. **Trust.** EarthCube design and governance should focus on fostering trust of data, and governance framework must create a broad and shared chain of trust from data providers to scientists and to end users of scientific results.

Topical and Case Study Recommendations

1. **AGORA-net.** AGORA-net graphically organizes knowledge based on argument maps that link arguments to one another and provide forums for debate, comment, and counterargument. The governance framework should be developed by the EarthCube community itself, and that governance must be open, transparent, and inclusive.
2. **Centralized Database.** EarthCube should be an "open access archive of all publicly available geosciences data in the United States and eventually the world,"[405] whose long-term vision would be led by scientists.
3. **CZO Data.** CZO Data's focus on efficient data and knowledge integration to understand and model earth processes is relevant to EarthCube.
4. **CUAHSI-HS.** CUAHSI-HS goals of data access, hydrologic observatories, hydrologic science, and hydrologic education, align well with EarthCube. Governance should focus on community process and open development. Specific mechanisms include establishing and documenting a governance model for development activities, creating tools to support community functions and interactions, and active engagement of partners and contributors.
5. **DataONE.** DataONE is an existing platform that allows scientists to focus on large-scale, long-term problems. DataONE employs stakeholder assessment, user scenarios, usability testing, and engagement with domain scientists in use cases to evaluate CI projects.

[405] Hannes Leetaru, "EarthCube – Strategic Organizational Framework (Centralized Model)," 1.

6. **EC Network.** EarthCube should be an ECNetwork of interconnected nodes, much like the internet, to build connections among data, resources, and disciplines. The governance framework should support distributed administration and available resources, create flexibility, promote scientific creativity, and take advantage of local expertise.
7. **EarthCube as a Double Loop Organization.** In Double Loop organizations, the processes and mechanisms developed to fix certain problems simultaneously solve larger issues that first created these problems. The governance framework allows stakeholders/community members to govern their own governance framework by agreeing to rules that solve common concerns related to interaction.
8. **Federation of Earth Science Information Partners (ESIP).** ESIP's community-based governance model fosters collaboration and coordination of interoperability efforts across the geosciences community and is a potential model for EarthCube.
9. **ESIP Discovery Cluster.** The ESIP Discovery Cluster is a volunteer-based organization that is consensus driven. The ESIP Discovery Cluster depends on open source software and collaborates to solve problems of interoperability, distributed services, and adoption of different open standards.
10. **Incorporated Research Institutions for Seismology (IRIS).** IRIS's community governance model provides a framework for decision-making that preserves community oversight of IRIS facilities and is a potential model for EarthCube governance framework. IRIS Unifying structure encourages operational efficiencies and coordination across the geosciences and is a good model for EarthCube governance.
11. **Open Geospatial Consortium (OGC).** OGC governance mechanisms include "a proven framework of tools, policies, and procedures for governance of long-term, interdisciplinary and multi-sector collaboration for high-impact scientific research for societal benefit."
12. **Professional Networking Service.** EarthCube should function as a professional networking service to link scientists to data, with a focus on professional relationships, protocols, formats, services and analytical tools. Governance should be community-based.
13. **QuakeSim.** The earthquake science field, and QuakeSim in particular, can be a potential partner to help provide the CI envisioned by EarthCube. QuakeSim promotes open processes, with a focus on the creation of data products.
14. **Scientific Forecasts.** EarthCube should focus on scientific forecasts. Governance would be divided into three groups: Economic, Scientific, and Technical.
15. **Space-Time Location.** A space-time location can serve as common ground for communication among earth scientists. EarthCube governance framework must focus on challenges in data integration, foster sharing and interdisciplinary collaboration, and maintain geophysical science information for the future.
16. **UNAVCO.** UNAVCO has created an effective community governance framework through a formal process of strategic planning and community involvement, which supports broad participation and effective oversight of UNAVCO programs.
17. **Unidata.** Unidata's governance framework focuses on assessing and responding to community needs, including advocating and negotiating agreements on behalf of the community on data issues: fostering community interaction to promote sharing of data, tools, and ideas; and consensus building.
18. **USGIN.** EarthCube should take advantage of the existing U.S. Geoscience Information Network (USGIN) and its tools and applications to understand data.
19. **USGS.** Collaboration among government entities and legislative mandates form bases of interaction between government data agencies and EarthCube participants. The USGS is already involved in building and sustaining cyberarchitecture to support its earth science mission.
20. **WInSAR Consortium.** The WInSAR governance model is an example of a low-cost and effective way to achieve both EarthCube and WInSAR goals.

CONCLUSION

The background research review on governance informs the entire Governance Roadmap-writing process and provides a key foundation on which to base Governance Roadmap recommendations and next steps. The key findings of the research review are summarized below.

KEY CONCEPTS FROM SECTION 1: INTRODUCTION TO GOVERNANCE

Governance Definitions:

- IT governance "specif[ies] the decision rights and accountability framework to encourage desirable behavior in using IT. IT governance is not about making specific IT decisions−management does that−but rather determines who systematically makes and contributes to those decisions."[406]
- Governance "provid[es] the structure for determining organizational objectives and monitoring performance to ensure that objectives are attained."[407]
- A "governance model describes the roles that project participants can take on and the process for decision making within the project. In addition, it describes the ground rules for participation in the project and the processes for communicating and sharing within the project team and community."[408]
- Governance "aligns an organization's practices and procedures with its goals, purposes, and values. Definitions vary, but in general governance involves overseeing, steering, and articulating organizational norms and processes (as opposed to managerial activities such as detailed planning and allocation of effort). Styles of governance range from authoritarian to communalist to anarchical, each with advantages and drawbacks."[409]

Types of Governance Reviewed:

1. Corporate IT governance
2. Non-profit governance
3. Open Source Software Governance

Effective governance must address three key questions:

1. What decisions must be made to ensure effective management and use of IT?
2. Who should make these decisions?
3. How will these decisions be made and monitored?

Five key types of decisions that must be made:

1. **IT principles**: Clarifying the business role of IT
2. **IT architecture**: Defining integration and standardization
3. **IT infrastructure**: Determining shared and enabling services
4. **Business application needs**: Specifying business need for purchased or internally developed IT applications
5. **IT Investment and prioritization**: Choosing which initiatives to fund and how much to spend

Governance archetypes that determine who makes decisions and how decisions are made:

---

[406] Weill and Ross, *IT Governance:* 2.
[407] Organization for Economic Cooperation and Development, 5, 219, quoted in Weill and Ross, *IT Governance*, 4-5
[408] Ross Gardler and Gabriel Hanganu, "Governance Models."
[409] "Governance," EarthSystem Commodity Governance Project, last modified 2012, http://earthsystemcog.org/projects/cog/governance_object

1. **Business Monarchy**: top managers
2. **IT Monarchy**: IT specialists
3. **Feudal**: each unit makes independent decisions
4. **Federal**: combination of corporate center and business units with or without IT people involved
5. **IT Duopoly**: IT group and one other group (ex: top management or business unit leaders)
6. **Anarchy**: isolated individual or small group decision making
7. **Meritocracy**: distributed control is awarded in recognition of contributions to the project
8. **Technical Utopianism**: technical approach to information management.

Governance Mechanisms: implement the governance framework on a day-to-day basis.

1. Decision-making structures
2. Alignment processes
3. Communication processes

Characteristics of Top Corporate Governance Performers

1. More managers in leadership positions can describe IT governance
2. Senior management is engaged in governance and communicates to the rest of the organization
3. More direct involvement of senior leaders in IT governance
4. Clear business objectives for IT investment
5. Differentiated business strategies
6. Fewer renegade decisions and more formally approved exceptions
7. Fewer changes in governance from year to year

Characteristics of Top Not-for-Profit Governance Performers

1. Executive committees focused on all key assets including IT
2. IT council comprising business and IT executives
3. IT leadership committee comprising IT executives
4. Architecture committee
5. Tracking of IT projects and resources consumed
6. Business/IT relationship managers

Top 10 leadership principles of effective governance:

1. Actively design governance
2. Know when to redesign
3. Involve senior managers
4. Make choices
5. Clarify the exception-handling process
6. Provide the right incentives
7. Assign ownership and accountability for IT governance
8. Design governance at multiple organizational levels
9. Provide transparency and education
10. Implement common mechanisms across the six key assets

A governance document for open source software projects should include:

1. Overview of project objectives, project management, identification of potential contributors and copyright holders, and licenses covering project outputs
2. Roles and responsibilities
3. Support processes within the project

4. Decision-making processes
5. Contribution process

KEY CONCEPTS FROM SECTION 2: HISTORICAL INFRASTRUCTURE DEVELOPMENT CASE STUDIES AND LESSONS LEARNED

Understanding Infrastructure: Dynamics, Tensions, and Design

Common steps in Historical Infrastructure Development

1. **System building:** Systems are linked into networks, leading to distributed control and access
2. **Technology Transfer:** Technology is transferred from one location or domain to another and the system adapts to new conditions
3. **Integration and Consolidation:** Heterogeneous systems are linked into networks through gateways, consolidating the infrastructure project

Infrastructure Dynamics

1. Reverse Salients
2. Gateways
3. Path Dependence
4. Scale Effects

Infrastructure Tensions

1. Time, Scale, and Agency
2. Winners and Losers
3. Interest and Exclusion
4. Ownership and Investment
5. Data Cultures and Data Tensions
6. Data Storage and Meaning
7. Data Sharing
8. Trust

Cyberinfrastructure Design Recommendations

1. **Purpose.** CI design should permit collaborative work that enables scientists from different domains to work together
2. **Boundary Work.** CI can bridge the 'great divide' between system building and social analysis
3. **Principles of Navigation.** It is necessary to establish a set of design principles of navigation instead of a rigid roadmap because there is no one way to design CI
4. **Standards and Flexibility.** Flexibility should be included in CI design to avoid path dependence and facilitate change

Recommendations for Future CI Projects

1. **Learn from successes and failures of past CI projects.** Comparing social scientific analysis of CI projects, using accurate and realistic reporting mechanisms, aligning extant or creating new incentive structures, and "instrumenting" CI projects for social scientific analysis will promote learning from past CI projects.
2. **Improve current CI analysis.** Infrastructural diagnostics, training for information managers, cross-disciplinary symposia for early-career scientists, and fitting funding to CI-development time scales can improve current CI analysis.
3. **Enhance CI resiliency and research.** CI resiliency, sustainability, and reach can be enhanced through design flexibility, mechanisms to incorporate under-represented groups and institutions, and alliances with niche organizations to eliminate redundancies in expertise.

Historical Case Studies

Project Mohole: Lessons Learned

1. Lack of governance structure led to failure of project.
2. Specific, unresolved issues that contributed to project failure included lack of clear purpose and objectives, leadership and a higher authority to resolve disputes, decision-making processes, communication mechanisms, and clear assignment of responsibility and accountability.
3. A governance framework could have created mechanisms to resolve these issues.
4. A governance framework needs to be put in place from the very beginning of a project in order to deal with anticipated and unanticipated challenges and issues as the occur.

Leadership Principles in NSFNET, Unidata, DLESE, NSDL

Key Leadership Principles

1. **Disciplined Thought:** Willingness to confront reality
2. **Disciplined Action:** Willingness to say "no"; clear identification of who holds responsibilities; advancement as a cumulative process
3. **Persistent Core:** An enduring purpose; immutable core values (around which to advance)
4. **Primary alignment with one of three value disciplines** (while maintaining the other two): Operational excellence; customer relations; product leadership

Leadership Recommendations

Create an **evaluation framework** with the involvement of leadership experts. An evaluation framework would eliminate redundancies and explain principles in the specific terms of the evaluation framework, thus avoiding informal and reflective analysis.

KEY CONCEPTS FROM SECTION 3: CURRENT STATE OF GOVERNANCE

The series of case studies demonstrates that governance is as varied as the institutions it serves, even within the same communities. **Case studies include:**

1. Domain science; CUAHSI, IRIS, Unidata
2. IT: Apache Software Foundation, DataNet, DataONE, W3C Consortium
3. Federations: ESIP and OGC
4. Large Facilities: GENI – Global Environment for Network Innovations, SURA US IOOS Coastal and Ocean Modeling Testbed

Case studies key points:

1. Each organization has a governance framework developed specifically to meet that organization's needs. For example, even within the federated organizations, OGC's governance framework differs greatly from that of ESIP.
2. Organizational goals vary greatly from organization to organization, as do the management structures and decision-making processes needed to achieve those goals.
3. Funding mechanisms and paid staff vs. volunteers differ greatly from organization to organization.
4. The effectiveness of the governance framework should be evaluated against how well the framework enables the organization to move forward in achieving its goals.

Geoinformatics Reports key points:

Earth and Space Informatics Summit

1. **A balance of leadership and governance** is required to successfully advance the geoinformatics agenda because leadership and governance have roles in how data and information efforts respond to user/stakeholder needs.
2. **Seven geoinformatics governance needs**:
    a. Chart (discover and map what's happening)
    b. Communicate
    c. Coordinate
    d. Contract (use expertise and products of others)
    e. Collaborate
    f. Converge (reduce complexity and proliferation)
    g. Seek Consensus

Workshop on Working towards a National Geoinformatics Community (NGC)

1. **Four major geoinformatics problems:**
    a. I can't integrate what I can't find
    b. I can't use something I don't understand
    c. I don't want to use something I don't trust
    d. I can't use something that isn't there anymore
2. **Community building** will address issues of data integration, sharing, longevity, and trust
3. **Recommends** a community-developed governance framework
4. **Recommends** hardware and software to promote technical innovation to advance the geoinformatics community

Conclusions from Geoinformatics Reports

1. There are actively engaged end-users in the geoinformatics community.
2. Best practices and lessons learned from these reports and others offer a wealth of knowledge from which to draw recommendations for EarthCube governance.

KEY CONCEPTS FROM SECTION 4: EARTHCUBE GOVERNANCE RECOMMENDATIONS FROM RESEARCH REVIEW

EarthCube White Papers (submitted fall 2011, Governance category and Design Category) contain several recommendations for EarthCube governance.

Thematic Recommendations:

1. Clarity
2. Collaboration
3. Commodity Governance and Communities of Communities
4. Community
5. Communities of Interest
6. Communities of Practice: Informaticists, Data Generators, Synthesizers
7. Community-based governance
8. Diverse Representation in Governance Framework
9. Evolution
10. Form Should Follow Function
11. Goals
12. Leadership
13. Open Source Software Governance and Open Science
14. Public-Private Partnership
15. Success
16. Sustainability
17. Trust

Topical and Case Study Recommendations

1. AGORA-net
2. Centralized Database
3. CZO Data.
4. CUAHSI-HS
5. DataONE
6. EarthCube as a Double Loop Organization (Virtual Democracy Blog recommendation)
7. EC Network
8. Federation of Earth Science Information Partners (ESIP)
9. ESIP Discovery Cluster
10. Incorporated Research Institutions for Seismology (IRIS)
11. Open Geospatial Consortium (OGC)
12. Professional Networking Service
13. QuakeSim
14. Scientific Forecasts
15. Space-Time Location
16. UNAVCO
17. Unidata
18. USGIN
19. USGS
20. WInSAR Consortium

APPENDIX 2: COMMUNITY ENGAGEMENT

The full potential of EarthCube capabilities will not be realized if community members are not aware of, or do not understand, tools and resources available.[410] Ongoing education and training, in addition to general community engagement, are essential to EarthCube success.[411] Conversely, a lack of education and outreach could lead to EarthCube failure, or at the very least, reduce EarthCube's effectiveness due to lack of training and education of the community for whom EarthCube will be created.[412] The importance of community engagement is clearly reinforced in the 44 out of 111 white papers submitted in the fall of 2011 that mention education and training,[413] in addition to several others that mention community building and engagement.

This appendix reviews materials from the background governance research review that focus on community engagement, outreach, training, education, and community-building mechanisms.

WORKFORCE (EDUCATION AND TRAINING) – END-USER ENGAGEMENT

Several EarthCube Expressions of Interest and White Papers focused on education and outreach, and are summarized here. Workforce development and education recommendations from the National Science Foundation Advisory Committee for Cyberinfrastructure Task Force on Cyberlearning and Workforce Development Final Report (March 2011) are also included, because they directly relate to the EarthCube specific education and outreach recommendations from the EoIs and white paper.

DATA MANAGEMENT TRAINING AND WORKFORCE DEVELOPMENT FOR THE EARTH SCIENCE COMMUNITY

Columbia CIESIN (1103): EOI purpose is to develop "New Capabilities that leverage existing technologies and practices and would significantly improve the productivity and capabilities of researchers and educators across all of the geosciences and that can be directly connected to the EarthCube vision."[414] The scientific motivation for the proposed effort is based on the following:

1. EarthCube will be a system of systems
2. Systems and networks of systems will likely greatly change how Earth science research is conducted
3. Creation of data management and informatics capabilities is a key component of EarthCube
4. The full potential of EarthCube capabilities will not be realized if community members are not aware of, or do not understand, tools and resources available.

Result and Advances that would be Enabled by Proposed Activities include:

1. Diverse community will have access to new capabilities to convey information about data management and informatics methods.
2. Targeted tutorials as needed on micro-topics related to data management and informatics
3. Greater understanding by community about importance and application of data management and new informatics methods

[410] Robert Downs, Robert Chen, and Mark Becker, "Data Management Training and Workforce Development for the Earth Science Community," National Science Foundation EarthCube Expression of Interest 1103, 2012, 1.
[411] Chaitanya Baru, "GET 21: Geoinformatics Education and Training for the 21st Century Geoscience Workforce," National Science Foundation EarthCube Expression of Interest 1054, 2012, 1, and Carlson, David, Shelley Olds, Steve Whitmeyer, Bruce Caron, Cathy Manduca, and John Tabor, "A Need for an Education and Outreach Component of EarthCube." National Science Foundation EarthCube White Paper: Designs Category, 2011.
[412] Chaitanya Baru, "GET 21," 2012, 1, and Robert Downs et al., "Data Management Training and Workforce Development," 2012, 1.
[413] Ibid.
[414] Robert Downs et al., "Data Management Training and Workforce Development," 2012, 1.

4. Resources developed for educational use in which data management becomes integral part of Earth science education

Proposed activities will:

1. Be built on expertise of U.S. partners in undergrad and graduate teaching and professional development
2. Engage new U.S. and international partners through the creation of hands-on workshops, and educational and training resources to be used independently or as part of new curriculum
3. Explore online game technology by developing interactive learning modules that support independent learners and workshop participants
4. Leverage previous experiences training Earth scientists about data management and informatics by providing training modules to be included in interdisciplinary educational programs and workshops
5. Conduct data management training workshops that will be packed into presentations to larger audiences, based on interest and experience, which will then be evaluated, revised, and shared publicly to increase their adoption throughout the EarthCube community

GET 21: GEOINFORMATICS EDUCATION AND TRAINING FOR THE 21ST CENTURY GEOSCIENCE WORKFORCE

Chaitanya Baru (1054): Geoinformatics education and training will help define what EarthCube will be, as geoscientists begin to fully understand the capabilities and limits of geoinformatics-based approaches. As informed consumers, EarthCube stakeholders are much more likely to maximize EarthCube value. Thus, EoI proposes to develop:

1. Specific programs that can be implemented right away and gain momentum by summer 2012
2. Develop set of New Capabilities in geoinformatics education
3. Meet Critical Milestones in preparing community for the effective use of CI

The EarthCube education agenda should address a broad range of communities, including graduate and post-doc students, junior and senior faculty, undergraduates, government agencies, and industry.
Specific proposed projects include:

1. Intensive, 4-6 week summer institute to provide geoscientists with geoinformatics skills needed to participate in EarthCube
2. Development of standard curriculum to be used across different universities
3. A 'teach the teachers' program
4. Internships at established geoinformatics facilities
5. A 'broker' service to connect students and researchers to geoinformatics experts
6. Online education and training resources for asynchronous learning

A NEED FOR AN EDUCATION AND OUTREACH COMPONENT OF EARTHCUBE

The NSF EarthCube White Paper "A Need for an Education and Outreach Component of EarthCube," submitted by Carlson et al., argues that a strong education and outreach component is needed form the very beginning in order to achieve the EarthCube goal of creating "a knowledge management system and infrastructure that integrates all geosciences data in an open, transparent and inclusive manner."[415] The link between education and research is

[415] David Carlson, Shelley Olds, Steve Whitmeyer, Bruce Caron, Cathy Manduca, and John Tabor, "A Need for an Education and Outreach Component of EarthCube," National Science Foundation EarthCube White Paper: Designs Category, 2011.

already clear in many organizations and in NSF GEO documents, and educational and research missions can help meet the "urgent need for efficiently managed and easily accessible data and data products."[416]

EarthCube will enhance access to data for many researchers, however, education and training is needed to inform the EarthCube community of new IT products, create the need and acceptance of new services, and reach out to current and future users. Education friendly formats for data need to be incorporated into EarthCube at the early stages of development because educators require higher standards of accessibility, reliability, and usability of research data. Specific requirements include quick and easy access to data in and understandable format, reliable and simple tools that are persistent from semester to semester, publicly accessible services that run on everyday platforms.

Additional educational needs include expanded concepts and definitions of metadata to suit educators, a curriculum and corresponding materials designed around data,[417] and a geo-referenced digital clearinghouse of images and digital products. An outreach strategy within the geosciences and to the larger scientific and public audience can help convey positive messages to the public regarding EarthCube, and keep educational media in line with current research and methods.

NATIONAL SCIENCE FOUNDATION ADVISORY COMMITTEE FOR CYBERINFRASTRUCTURE TASK FORCE ON CYBERLEARNING AND WORKFORCE DEVELOPMENT FINAL REPORT, MARCH 2011

The EarthCube-specific education and outreach recommendations presented in the EoIs and White Paper described above align very well with the recent "National Science Foundation Advisory Committee for Cyberinfrastructure Task Force on Cyberlearning and Workforce Development Final Report" (March 2011). The report argues that there is a great need for science, technology, engineering, and mathematics (STEM) education, particularly in computational and data-intensive science and engineering (CDS&E), and the skillful use of CI for knowledge creation and learning.[418]

NSF can foster STEM education by investing in efforts to better understand learning and research mechanisms and organizations in an environment of rapid technological advances, partnering with other Federal Agencies, encouraging bold, revolutionary proposals, workforce development, K-12 training, informal science education, and lifelong learning, bridging college campuses into CI for Cyberlearning and workforce development, and broadening participation by focusing on eliminating underrepresentation of women, minorities, and people with disabilities in STEM fields and national scientific leadership positions. Specific workforce development and education recommendations include:

Workforce Development

1. Promotion and support of a cross-disciplinary community to perform transformative research and translate research for increasingly heterogeneous communities and end-users (focusing on needs of women, minorities, and people with disabilities).
2. Development of a systematic strategy to encourage cumulative outcomes from item 1.
3. Address two critical challenges of systematic change: 1) Develop models for educational system organization and processes that respond quickly and appropriately to changing technologies, and 2) Define skills and knowledge needed to support a strong nationally and internationally connected economy and workforce.

[416] Ibid.
[417] T.S. Ledley, A. Prakash, C.A. Manduca and S. Fox, "Recommendations for Making Geoscience Data Accessible and Usable in Education," EOS 89, no. 32, (2008): 291, quoted in Carlson et al., "A Need for an Education and Outreach Component of EarthCube," 2.
[418] National Science Foundation Advisory Committee for Cyberinfrastructure Task Force on Cyberlearning and Workforce Development, "Final Report," March 2011,

4. Creation of a new generation of interdisciplinary thinking by taking advantage of advancements in disciplines related to cyberlearning.

K-12, Training, Informal Science Education, Lifelong Learning

1. CI-enabled research, tools, resources, advancements should be refocused on lifelong learning and professional development
2. Lifelong learning challenges the NSF to create and support bridges among different governmental agencies, industry, education
3. Continual professional development of educators Resources for teachers to understand and adopt new pedagogies using new technology
4. Keep abreast of new developments and incorporate into teaching
5. Support research in cyberlearning
6. Explore meaningful metrics to assess needs and progress of all learners
7. Support research on methods to attract and retain diverse STEM workforce
8. Structure funding programs to support interdisciplinary teams working on architecture and implementation of cyberlearning platforms that are open, scalable, flexible, sustainable
9. Support for cyberlearning resources
10. Repurposing of research data, tools, resources for learning purposes
11. Access by diverse audiences and purposes
12. Promote modeling and simulation, use multiple representations, quantitative reasoning, parallel methods across disciplines and throughout lifelong learning process

COMMUNITY BUILDING & VIRTUAL COMMUNITIES

Identification of the EarthCube community and end-user engagement (education and training, mechanisms to encourage contributions and ownership of decision-making processes, etc.) will likely be key EarthCube governance challenges. Investigations focusing on these issues offer valuable insights on the challenges that will likely confront EarthCube. These studies emphasize specific mechanisms that promote community creation and maintenance, and include *How Oversight Improves Member Maintained Communities* by Cosley et al. (2005), *Encouraging Participation in Virtual Communities* by Broß, Sack and Meinel (2007), *Sense of Virtual Community—Maintaining the Experience of Belonging,* by Blanchard and Markus (2002), and *Harnessing Crowds: Mapping the Genome of Collective Intelligence* by Malone, Laubacher and Dellarocas (2009).

HOW OVERSIGHT IMPROVES MEMBER MAINTAINED COMMUNITIES

Cosley, Frankowski, Kielser, Terveen and Riedl argue that increased oversight of virtual communities can result in greater quantity and quality of community member contributions. Although online communities require regular maintenance activities, (such as moderation and data input), these activities are usually done by community owners. Communities that allow members to participate in maintenance are tasks usually better developed and more sustainable, however, than owner-communities. Member-maintained communities experience reduced reliance on key individuals, they draw on the resources of all community members, and scale as the community grows.

Why Do People Contribute?

Cosley et al. cite many theories as to why people contribute to communities (virtual and face-to-face). Social dilemma is one factor that affects individuals' contributions. It occurs when a group would benefit from making a certain choice, but members have incentives to make the opposite choice.[419] Contributions to community maintenance are a social dilemma when people decide not to contribute if others are contributing,[420] even though the community would benefit from more contributions. The Collective Effort Model[421] explains that the motivation for any work depends on how well that work translates into overall performance.[422] When deciding whether to contribute, people weigh likely outcomes of expected performance with the value of those expected outcomes, taking into account individual and group perspectives.

Experiment Results and Conclusions

To test the difference between owner and member-maintained communities in terms of quantity and quality of member contributions, Cosley et al. asked the research question, "How can we design mechanisms—interfaces, algorithms, economies, and social structures—that allow communities to maintain themselves and encourage members to provide valuable contributions?"[423] and focused on the website MovieLens to conduct their research.

Cosley et al. found that oversight improves overall outcomes of member-maintained communities and increases contributions. Specifically, oversight mechanisms improve the quality of the community and reduce antisocial behavior. Peer-versus-expert oversight had very little effect the quantity or quality of member contributions, and that the overall quality of community contributions improved by selecting for the best contributors and improving the capabilities of individual users through training. They also suggest informing members about oversight as a potential means increase motivation to contribute.

ENCOURAGING PARTICIPATION IN VIRTUAL COMMUNITIES: THE "IT-SUMMIT-BLOG" CASE

In contrast to Cosley et al., Broß, Sack, and Meinel in "Encouraging Participation in Virtual Communities: The "IT-summit-blog" argue that trust, recognition vs. anonymity, time, differing needs of community members, and sense of community are key factors in determine people's motivation to contribute, although different systems and bring about differing user needs. For example, anonymity encourages some people to contribute, while other individuals are motivated to contribute to communities that recognize and reward contributions, in the form of financial rewards, elevated social status[424] and visibility gained through participation in virtual communities.[425] A sense of community can also affect member participation, as do time available, level of interest,[426] and amount of trust community members feel towards others in the community.

[419] Robyn Dawes. "Social Dilemmas," *Annual Review of Psychology* 31 (1980): 169–193, cited in Cosley et al., "How Oversight Improves Member Maintained Communities,"2.

[420] Norbert Kerr, "Motivation Losses in Small Groups: a Social Dilemma Analysis," *Journal of Personality and Social Psychology*, 45 no. 4 (1983): 819–828, cited in Cosley et al., "How Oversight Improves Member Maintained Communities," 2.

[421] Steven Karau and Kipling Williams, "Social Loafing: a Meta-Analytic Review and Theoretical Integration," *Journal of Personality and Social Psychology*, 65 no. 4 (1993): 681–706, quoted in Cosley et al., "How Oversight Improves Member Maintained Communities," 2.

[422] Victor Vroom, *Work and Motivation*. (New York: Wiley, 1964,) cited in Cosley et al., "How Oversight Improves Member Maintained Communities," 2.

[423] Cosley et al., "How Oversight Improves Member Maintained Communities," 2.

[424] Dorine Andrews, "Audience-Specific Online Community Design," *Communications of the ACM,* 54 no. 4 (2002):64-68, Calvin Chan et al, "Recognition and Participation in a virtual community," *Proceedings of the 37th Hawaii International Conference on System Sciences* (2004); John Hagel and Arthur Armstrong, *Net Gain: Expanding Markets through Virtual Communities,* (Boston: Harvard Business School Press, 1997), cited in Justus Broß, Harald Sack, and Christoph Meinel, "Encouraging Participation in Virtual Communities: The "IT-summit-blog" Case," *IADIS International Journal on WWWInternet*, 5 no. 2 (2007), 4.

[425] Brian Butler, et al., "Community Effort in Online Groups: Who Does the Work and Why?" Human-Computer Interaction Institute, Paper 90, http://repository.cmu.edu/hcii/90; Brian Butler, "Membership Size, Communication Activity, and Sustainability: the Internal Dynamics of Networked Social Structures," *Information Systems Research*, 12, no. 4 (2001): 346–362, cited in Broß et al., "Encouraging Participation in Virtual Communities," 5.

[426] McClure Wasko, and Samer Faraj "It is What One Does: Why People Participate and Help Others in Electronic Communities of Practice," *Journal of Strategic Information Systems*, 9, no. 2-3 (2000): 155-173, cited in Broß et al., "Encouraging Participation in Virtual Communities," 4.

Broß et al. cite four incentives to promote participation: leadership, offline interaction, usefulness, and IT-infrastructure quality.[427] Leadership involves building relationships and developing user-created content,[428] in addition to encouraging other community members to actively post or view community content, and promoting collaboration, trust, and clear objectives. Offline interaction can have a positive influence on the community by promoting understanding, trust, and community member awareness of other members. Usefulness refers to the benefits of membership surpassing the costs,[429] meaning that community members view the content as valuable, interesting, and useful. [430] IT infrastructure should promote efficient communication, and awareness of other people and interpersonal relationships during an interaction.

SENSE OF VIRTUAL COMMUNITY—MAINTAINING THE EXPERIENCE OF BELONGING

Blanchard and Markus argue that a sense of community (SOC) in virtual communities is similar to SOC in face-to-face communities, but with a few key differences. In their research on the MSN virtual community, they sought to answer the following questions:[431]

1. "Does a sense of community similar to that sometimes observed in physical communities also occur in virtual settings?"
2. "How is a sense of virtual community maintained?"

Sense of Community Theory

Blanchard and Markus base their research on SOC, defined as, "a feeling that members have of belonging, a feeling that members matter to one another and to the group, and a shared faith the members' needs will be met through their commitment to be together."[432] SOC can lead to important outcomes in face-to-face communities and organizations. In the work environment these outcomes include increased job satisfaction, and more organizational citizenship behavior (Loyalty, civic virtue, altruism, courtesy[433]). SOC can be the outcome of living in a community and can be the definitions of the community itself,[434] although SOC is harder to define for communities that are not place-based (such as virtual communities). SOC may be unique to each community, and may include one or all of the following components: feelings of membership, feelings of influence, integration and fulfillment of needs, and shared emotional connection. [435]

Blanchard and Markus' found that SOC arises in virtual communities from social processes that maintain the SOC. They found that SOC in the MSN community was maintained through mutual exchange of support, creation of identities, making of identifications, and production of trust. Specifically, they highlight membership, influence, and intimacy as three basic group dynamics in virtual communities,[436] and identify three interrelated social

---

[427] Joon Koh, J. and Young-Gul Kim, "Sense of Virtual Community: Determinants and the Moderating Role of the Virtual Community Origin," *Proceedings of the 22nd International Conference on Information Systems*, (2001): 407-409 cited in Broß et al., "Encouraging Participation in Virtual Communities," 4.

[428] Cliff Figallo, *Internet World- Hosting Web Communities: Building Relationships, Increasing Customer Loyalty, and Maintaining a Competitive Edge,* (New York: John Wiley & Sons, Inc., 1988), cited in Broß et al., "Encouraging Participation in Virtual Communities," 4.

[429] Brian Butler, et al., "Community Effort in Online Groups: Who Does the Work and Why?" and Brian Butler, "Membership Size, Communication Activity, and Sustainability," cited in Broß et al., "Encouraging Participation in Virtual Communities," 5.

[430] John Hagel and Arthur Armstrong, *Net Gain,* cited in Broß et al., "Encouraging Participation in Virtual Communities," 4.

[431] Anita Blanchard and Lynne Markus, "Sense of Virtual Community-Maintaining the Experience of Belonging," *Proceedings of the 35th Hawaii International Conference on System Sciences, Computer Society, Institute of Electrical and Electronics Engineers*, 2002, 1.

[432] David McMillian and David Chavis, "Sense of Community: A Definition and Theory," *Journal of Community Psychology*, 14, (1986): 6-23, cited in Blanchard and Markus "Sense of Virtual Community," 2.

[433] Susan Burroughs and Lillian Eby,"Psychological Sense of Community at Work: a Measurement System and Explanatory Framework," *Journal of Community Psychology*, 26, (1998): 509-532, cited in Blanchard and Markus "Sense of Virtual Community," 1.

[434] Isabel Garcia, Fernando, Giuliani and Esther Wiesenfeld, "Community and Sense of Community: the Case of an Urban Barrio in Caracas", *Journal of Community Psychology,* 27, (1999): 727-740, cited in Blanchard and Markus "Sense of Virtual Community," 2.

[435] Ibid.

[436] Wilfred Bion, *Experiences in Groups*, (London: Tavistock, 1961), cited in Blanchard and Markus "Sense of Virtual Community," 8.

processes: exchange of support, production of trust, and creation of identify (distinguishing oneself from the group) and making of identifications.

In conclusion, Blanchard and Markus argue that SOC is much more than 'build it and they will come.' Virtual community creators/owners must pay attention to the type of virtual community they want to create, because members will only participate if they see a benefit. Thus, "a virtual community is a virtual settlement in which a sense of virtual community has emerged from a set of community-like behaviors and processes. Community-like processes and the sense of virtual community outcome cannot be guaranteed. They require people to enact them and to continue enacting them over time. Thus, understanding how such processes get started should remain high on the agenda for research in the virtual communities tradition."[437]

HARNESSING CROWDS: MAPPING THE GENOME OF COLLECTIVE INTELLIGENCE

Malone, Laubacher and Dellarocas in "Harnessing Crowds: Mapping the Genome of Collective Intelligence" identify the building blocks of collective intelligence that can help determine Who, What, Why and How in the decision-making process. Malone et al. define collective intelligence as, "Groups of individuals doing things collectively that seem intelligent,"[438] and present Google, Wikipedia, and Threadless as examples of modern, internet-enabled forms of collective intelligence. Organizational building blocks are the core of collective intelligence systems, and there are conditions under which each block is useful. In addition, there are endless possibilities to combine and re-combine these genes to effectively harness crowds, depending on the project goals, structure, and desired outcomes.

Collective intelligence building blocks (divided into organizational 'genes' by the authors) determine who is performing a specific task, why they are doing it, answer what is the goal of the task, and how it will be completed. Specifically, they identify the necessary staffing (who), goals/objectives (what), incentives (why) and structure/process (how). A 'gene' is a specific answer to *Who? What? Why?* and *How?*, and a genome is a combination of genes related to a specific example of collective intelligence.

- *Who?* Asks if authority is hierarchical (an authority figure assigns specific tasks for the group), or crowd-sourced (anyone can undertake a task – crowd sourcing is a key component of any internet based collective intelligence).
- *Why?* Asks why people participate, and what are their motivation and/or incentives. Typical incentives include money, love (enjoyment of the activity) and recognition.
- *What?* Asks what is being done. The answer can be the mission statement or goal.
- *How?* Asks how the collective intelligence system will be structured and what processes it will employ. It determines how products will be created (via independent contributions or collaborations, or some combination of the two), and how decisions will be made. It determines if decision-making will be a group or individual process.

Malone et al. found that crowd-sourcing is commonly used for creation and intermediate decisions, although final decisions are usually made by a smaller group of individuals assigned to the task. They recommend crowd-sourcing when, "The resources and skills needed to perform an activity are distributed widely or reside in places that are not known in advance,"[439] and it is possible to divide the activity into discrete pieces that can be accomplished by different people. If crowd-sourcing is to be used, however, the authors recommend mechanisms to prevent cheating or sabotaging the system.

---

437 Blanchard and Markus "Sense of Virtual Community," 8.
438 Malone, Thomas, Robert Laubacher and Chrysanthos Dellarocas, "Harnessing Crowds: Mapping the Genome of Collective Intelligence," MIT Center for Collective Intelligence, 2009, 1.
439 Ibid., 15

Malone et al. argue that it is very important to choose the right motivations to get people to contribute. Focusing on love or recognition may reduce costs, but these motivations don't always work. They suggest that money and recognition can move the process along and enable more rapid achievement of specific goals.

As to whether decisions should be group or individual, Malone et al. suggest that group decisions are best when everyone in the group is bound by the same decision. Consensus is a good decision-making process for a small group, but voting is better for larger groups and when everyone must be committed to the outcome. In contrast, individual decisions are useful when money or other incentives are needed to motivate people to contribute effort and/or resources.

In conclusion, Malone et al. suggest that it is important to evaluate each activity to be done and consider the many possible combinations of Who, What, Why, and How.

FINAL THOUGHTS ON VIRTUAL COMMUNITIES

A common conclusion among almost all the research reviewed in this section is that the 'build it and they will come' theory is not viable, and that if a robust virtual community is to thrive and be sustainable, active community building mechanisms must be in place from the very beginning. Community building mechanisms include, but are not limited to, active leadership, oversight, creation and maintenance of a sense of community, incentives to contribute, and active planning to decide Who, What, Why and How a specific task will be accomplished. EarthCube will most likely be a primarily virtual community, composed of multiple (and sometimes overlapping) communities of practice distributed across the United State and potentially world. Thus, the findings highlighted by these authors on virtual community-building will be very pertinent to the eventual EarthCube governance framework.

APPENDIX 3: PROCESS (EXPANDED)

The EarthCube community is fractal – projects from individual investigators, to teams or groups, to institutions and domain-wide or geographically distributed communities. Each Governance Steering Committee member's activities connect directly from 10s to 100s of others, although they range widely in their cyberinfrastructure engagement or interest. Each of the Steering Committee members is also active in multiple cyberinfrastructure activities, again, providing established conduits to readily and quickly connect with the broad community. The makeup of the Steering Committee and Forum covers multiple players from academia, government (both state and federal), and the private sector, domestically and internationally. These diverse characteristics helped ensure that we connect with the full breadth of the community and those who wish to join the community.

TEAM MEMBERS

Members of the ad-hoc EarthCube Governance Working Group Steering Committee were drawn initially from the teams that submitted Expressions of Interest to the NSF EarthCube initiative in January 2012 that addressed community consensus process and standards governance mechanisms. Members of Steering Committee, their organizational affiliations and cyberinfrastructure affiliations, and their EarthCube Expression of Interest affiliation or white paper are listed in Table 21 below.

EarthCube Governance Steering Committee Members

Name	Institution	Expression of Interest
Allison, M. Lee: PI	Arizona Geological Survey (AZGS), [US Geoscience Information Network, National Geothermal Data System, GeoNet Federation],	1064
Crane, Gary: Co-PI	Southeastern Universities Research Association (SURA), [SURAGrid, US IOOS Coastal Ocean Modeling Testbed]	1016
Arctur, David: Co-PI	Open Geospatial Consortium (OGC)	1053
Leetaru, Hannes: Co-PI	Illinois Geological Survey, University of Illinois	1015
Fox, Geoffrey: Co-PI	Indiana University [QuakeSim]	1055
Zaslavsky, Ilya	San Diego Supercomputer Center (SDSC)	1051
Ramamurthy, Mohan	UCAR [Unidata]	1053/white paper
Robinson, Erin	Federation for Earth Science Information Partners, ESIP	1064
Ahern, Tim	IRIS Data Management	1041
Lehnert, Kerstin	Columbia University, Lamont Doherty Earth Observatory , [IGSN, EarthChem, Geoinformatics for Geochemistry, Integrated Earth Data Applications, PetDB, SedDB, SESAR]	1041 & 1053
Bowring, Jim	College of Charleston	1022

TABLE 21: GOVERNANCE WORKING GROUP STEERING COMMITTEE MEMBERS

Initial EoIs included in the Steering Committee are listed below (and summarized in the next section of this document).

- Strategic Organizational Framework (Centralized Model) – Illinois (1015)

- Designing a community-based framework for modeling the consequences of climate change on coastal systems – SURA (1016)
- An Open Geoscience Community – CUAHSI (1019)
- Organizational Framework for EarthCube – OGC (1053)
- OpenQuake Social Networking for EarthCube – USC/JPL (1055)
- GeoNet: A Confederation of Scientific Data Networks – AZGS (1064)
- Federation of Earth Science Information Partners (ESIP) *- Many ESIP members are leveraging this open collaboration platform and flexible governance system for EarthCube activities as referenced in EoIs 1014, 1029, 1061 and 1084 among others.*

We reviewed the other EoI's submitted to NSF that were aggregated under the three other Community Groups ('Blue teams") and the Concept Award Teams ("Amber teams") and found three EoIs in particular with strong governance components to them. Because of this, the lead authors of each of those EoIs were added to the EC-Gov Steering Committee.

- Readiness of Disciplinary Data Systems for Cross-Domain Interoperability within a Standards-Based EarthCube Reference Framework – San Diego Supercomputer Center (1051)
- Service Based Vertical and Horizontal Integration Platform – IRIS Data Management Center (1041)
- Building a Software Engineering and Informatics Community of Practice and a Software Foundation for the Advancement of EarthCube – College of Charleston (1022)

EARTHCUBE GOVERNANCE FORUM

We sought representatives from each of the partners on the Steering Committee's EoIs to form a larger, broader constituency that had demonstrated their interest and participation in governance issues and from each of the EarthCube Community Groups ("Blue" groups) and Concept Award Teams ("Amber" groups). The latter served as liaisons between the more technically focused groups and the Governance Working Group to help ensure unique governance issues and needs from each segment of EarthCube are identified and addressed.

One of the goals of organizing the Forum was to ensure representation from each of the scientific domains (e.g. Earth, Atmospheres, Ocean, Environment, Ecology), IT sectors, academia, agencies, and industry, involved in determining agendas, setting goals, and having an active role in formulating the roadmap. Below is a list of the organizations who submitted letters of support, expressing their interest in collaborating with the Governance Steering Committee.

A list of organizations who submitted letters of support and interest in collaborating with the Governance Steering Committee is included in Table 22 below:

EarthCube Governance Forum Members

	Organization name	EoI Lead	EoI Number
1	AZGS	GeoNET	1064
2	CIRA, Colorado State University	College of Charleston	1022
3	Clemson University	OGC	1053
4	College of Charleston / EARTHTIME / CIRDLES	College of Charleston	1022
5	Columbia University CIESIN	OGC	1053
6	Columbia University IEDA	IRIS & OGC	1041 & 1053
7	CUAHSI	IRIS	1041 & 1053
8	Earthscope	OGC	1053
9	ESIP Federation	GeoNET	1064
10	Geoscience Australia	GeoNET	1053 & 1064
11	Illinois Geological Survey	University of Illinois	1015
12	Indiana University	Indiana University - QuakeSim	1055
13	IRIS Data Management	GeoNET	1041
14	Italian National Research Council	OGC	1053
15	NASA Jet Propulsion Laboratory	Indiana University - QuakeSim	1055
16	Johns Hopkins Data Conservancy	OGC	1053
17	National Center for Supercomputing Applications	University of Illinois & College of Charleston	1015 & 1022
18	National Ecological Observatory Network	GeoNET	1064
19	National Snow and Ice Data Center	University of Illinois	1015 & 1053
20	OGC	OGC	1053 & 1064
21	OPeNDAP, Inc.	IRIS	1041
22	RENCI, Univ. of NC Chapel Hill	College of Charleston	1022
23	San Diego Supercomputer Center	OGC	1053
24	SURA	SURA	1016
25	Texas Advanced Computing Center/iPlant Collaborative	GeoNET	1064
26	UNAVCO	IRIS	1041
27	UNIDATA	UCAR and OGC	1053
28	UC Davis	Indiana University - QuakeSim	1055
29	UCAR	UCAR and OGC	1053
30	UCSD/Calit2/Scripps OOI CI	OGC	1053
31	University of Illinois	Univ. of Illinois	1015
32	University of Miami, Center for Computational Science	College of Charleston	1022
33	USC	Indiana University - QuakeSim	1055
34	USGIN	GeoNET	1053 & 1064

TABLE 22: LETTERS OF SUPPORT RECEIVED BY EARTHCUBE FORUM MEMBERS FOR EARHTCUBE GOVERNANCE WORKING GROUP

EARTHCUBE GOVERNANCE WORKING GROUP AND CONCEPT TEAM LIAISONS

In addition to the EarthCube Governance Steering Committee and Forum, we also identified a representative from each EarthCube Working Group and Concept Team to act as a liaison between their group and the Governance Working Group (Table 23 & 24)). Liaisons served a key role in providing feedback on their governance needs and wants.

Concept Team	Liaison to Governance Group	Principal Investigator
Data Brokering	Ruth Duerr,	Jay Pearlman
Earth System Models	David Gochis, Richard Hooper, Scott Peckman, Jennifer Arrigo	David Gochis
Layered Architecture	Reagan Moore	Reagan Moore
Cross-Domain Interoperability	David Arctur, OGC	Ilya Zaslavsky
Web Services	Tim Ahern	Tim Ahern

TABLE 23: LIAISONS TO THE GOVERNANCE WORKING GROUP FROM CONCEPT TEAMS

WORKING GROUPS

Working Group	Representative to Governance	Organization
Data Discovery, Access and Mining	Jo Baker Lee Allison Chaitan Baru	Virginia Tech Arizona Geological Survey San Diego SuperComputer Center
Semantics and Ontologies	Andrew Maffei	Woods Hole
Workflow	Yolanda Gil	Information Sciences Institute

TABLE 24: LIAISONS TO THE GOVERNANCE WORKING GROUP FROM THE OTHER WORKING GROUPS

STEERING COMMITTEE EOIS

The EC-Gov Steering Committee was drawn initially from the teams that submitted Expressions of Interest (EoIs) to the NSF EarthCube initiative in January 2012 that addressed community consensus process and standards governance mechanisms. Reviewing these EoIs was the starting point for conducting a series of workshops on EarthCube. We summarize key elements of each EoI below (title of EoI, submitting organization, and EarthCube number):

STRATEGIC ORGANIZATIONAL FRAMEWORK (CENTRALIZED MODEL)

Illinois Geological Survey (1015): – centralized open access archive of all publicly available geosciences data in the U.S. and eventually the world.[440]

The centralized model proposed here is explicitly focused on long-term sustainability and access, using an approach similar to the GenBank and NSF LSST projects. Under the EAGER award, the feasibility of several funding models will be explored, including annual per-disk use and membership fees, to test sustainability.

[440] Leetaru, Hannes, "EarthCube – Strategic Organizational Framework (Centralized Model)." National Science Foundation EarthCube Expression of Interest 1015, 2012.

DESIGNING A COMMUNITY-BASED FRAMEWORK FOR MODELING THE CONSEQUENCES OF CLIMATE CHANGE ON COASTAL SYSTEMS

SURA (1016): The overarching goal of the envisioned program is to create a multi-disciplinary, community modeling cyber environment that will improve predictability and understanding of natural and anthropogenic influences on coastal ecosystems.[441]

We advocate the development of an enduring, comprehensive cyberinfrastructure that will enable the assessment and validation of models designed to advance our understanding of the physical oceanographic, geological, biogeochemical, fluvial hydrological processes and human-induced modifications that interact to cause environmental and ecological changes that impact humans in profound ways. The infrastructure must provide essential data on all aspects of the problem and must enable coupled modeling of coastal processes, including socioeconomic processes. The new infrastructure must enable the modeling of process suites that may be forced or influenced by potential climate change scenarios that have already been predicted by modelers at NCAR and elsewhere.

Tools that leverage or, as necessary, define community standards to enable the efficient access, visualization, skill assessment and other evaluations of multiple model results as well as tools that enable the coupling of multiple models

A cyber-environment where researchers can work together on selected modeling applications as a means of fostering and enhancing the development of new and improved coastal systems models that is scalable to multiple geoscience domains.

…support a series of community development, planning and design workshops that will bring together scientific stakeholders, model developers and cyberinfrastructure architects to design a working cyberinfrastructure to support the development of coastal systems models that is extensible to the multi-disciplinary science requirements of a Phase 1 Earth Cube. Following the workshops, a report will be prepared that details a comprehensive research plan for the future and the science requirements of a cyberinfrastructure architecture to support the envisioned research. An underlying theme of the effort will be: Connecting models across multiple disciplines, building the required cyber and data infrastructure and assessing the skill of models.

AN OPEN GEOSCIENCE COMMUNITY

CUAHSI (1019): A very large challenge for EarthCube is to build consensus among this diffuse community about common standards and approaches that are valuable for discovering, accessing, and processing existing data, and publishing new datasets and models that support enhanced information discovery and knowledge development.[442]

It would be relatively easy to hold a parallel meeting [with OGC's Technical Committee] of academic representatives of the EarthCube community to consider the challenges that EarthCube presents for information sharing in the academic geoscience community, and seeks to learn from the experience over the past fifteen years that the Open Geospatial Consortium has developed on how to build consensus towards community standards for information interchange. This discussion needs to acknowledge the particular characteristics of the academic

---

[441] Crane, Gary, "Designing a Community-Based Framework for Modeling the Consequences of Climate Change on Coastal Systems," National Science Foundation EarthCube Expression of Interest 1016, 2012.
[442] Maidment, David, "EarthCube – An Open Geoscience Community." National Science Foundation EarthCube Expression of Interest 1019, 2012.

geoscience community, namely that it is a loosely connected group of individuals who each have their own perspective, and who seek to differentiate themselves from their peers by developing a body of knowledge and expertise that is personal to the individual, or to a small science team that works together, such as with a professor and associated group of graduate students. It is a characteristic of such a knowledge development process that the researchers strongly "own" the way that they process their data and develop their conclusions. A very large challenge for EarthCube is to build consensus among this diffuse community about common standards and approaches that are valuable for discovering, accessing, and processing existing data, and publishing new datasets and models that support enhanced information discovery and knowledge development.

ORGANIZATIONAL FRAMEWORK FOR EARTHCUBE

OGC (1053): OGC seeks to lead and facilitate a consensus-based community process to advance the study of, and planning for, a Strategic Organizational Framework for EarthCube. During OGC's 18-year history, we have evolved consensus-based practices and policies for standards development to enable and improve the interoperability of data and processes across disparate global user communities.[443]

The main elements of a strategic organizational framework for the EarthCube Community include:

1. Regularly occurring, face-to-face meetings of the EarthCube Community.
2. Sustained presence of a virtual community between EarthCube meetings.
3. Definition of consensus policy and procedures for the EarthCube Community.
4. Establishment of working groups for geoscience and cyberinfrastructure topics.
5. Promotion of incubator projects by NSF funded activities that inform the development of EarthCube.
6. Consensus agreement on a reference architecture and baseline of open standards required for use in the EarthCube cyberinfrastructure.

OPENQUAKE SOCIAL NETWORKING FOR EARTHCUBE

USC/JPL (1055): explore the development of social networking and information management tools to improve the ability of earthquake researchers to discover and share data and undertake collaborative research.

We believe that existing social networks (particularly Ning (as used in EarthCube), Google+ and Facebook) can be leveraged to create an overlay social network that can deliver more specialized services and effective collaboration. Our team's e-Humanities effort (http://ehumanity.org/) illustrates some of these ideas for a different domain. We will also look at lessons from the well-known NanoHub portal.

GEONET: A CONFEDERATION OF SCIENTIFIC DATA NETWORKS

AZGS (1064): GeoNet is a scientifically broad-based confederation of scientific data networks, principally in the geosciences, working to build a community of practice (CoP) converging towards the creation of an integrated global digital data framework.[444] We plan to leverage existing and emerging technologies and practices in an effort to facilitate enhanced and interoperable cyberinfrastructure. In leveraging technologies, organizations, and procedures in their incubation stages, the CoP will also be able to provide input and recommendations on achieving critical EarthCube capabilities in best practices, definitions, and standards on data management.

---

[443] Arctur, David, "Expression of Interest in Development of an Organizational Framework for EarthCube." National Science Foundation EarthCube Expression of Interest 1053, 2012.
[444] Allison, Lee, "GeoNet: A Confederation of Scientific Data Network." National Science Foundation EarthCube Expression of Interest 1064, 2012.

FEDERATION OF EARTH SCIENCE INFORMATION PARTNERS (ESIP)

The ESIP Federation is a distributed community of science, data and information technology practitioners that coordinate interoperability efforts across the Earth and environmental science communities. Receiving support from multiple federal agencies, the ESIP Federation realizes its mission: "*To support the networking and data dissemination needs of our members and the global community by linking the functional sectors of observation, research, application, education and ultimate use of Earth science"* by proving a forum (virtual and in person) for exchanging knowledge and working collaboratively toward goals benefitting interests broader than any single partner. Together, the ESIP Federation, the community it represents, and its sponsors have tackled challenges facing the environmental data and information community for over 14 years while working toward improved data and information delivery for public benefit. Many ESIP members are leveraging this open collaboration platform and flexible governance system for EarthCube activities as referenced in EoIs 1014, 1029, 1061 and 1084 among others.

ADDITIONAL GOVERNANCE VISIONS FROM EARTHCUBE PARTICIPANTS (AND SC MEMBERS)

We reviewed the other EoI's submitted to NSF that were aggregated under the three other Community Groups ('Blue teams") and the Concept Award Teams ("Amber teams") and found three EoIs in particular with strong governance components to them. Because of this, the lead authors of each of those EoIs were added to the EC-Gov Steering Committee.

READINESS OF DISCIPLINARY DATA SYSTEMS FOR CROSS-DOMAIN INTEROPERABILITY WITHIN A STANDARDS-BASED EARTHCUBE REFERENCE FRAMEWORK

San Diego Supercomputer Center (1051): This project will systematically assess community readiness for EarthCube interoperability framework with respect to the four key formal domain infrastructure components and use this as the basis for organizing a community consensus process and standards governance mechanisms.[445]

SERVICE BASED VERTICAL AND HORIZONTAL INTEGRATION PLATFORM

IRIS Data Management Center (1041): We believe a service-based architecture can be used to integrate data, information, and knowledge products both within and across earth science disciplines, consistent with the EarthCube vision.[446] In fact we believe loosely coupled web services offer the best technical approach to an integrated EarthCube that exhibits the flexibility needed to evolve. The team endorses and will follow the Guiding Principles embraced within Amazon including:

- All disciplines will expose their data, information and knowledge assets through service interfaces.
- All technical communication between and within disciplines must be through these service interfaces.
- There will be no other interprocess communication allowed. All developments must make use of exposed services and not allow any backdoor access to information.
- All service interfaces, without exception, must be designed from the ground up to be exposed to external developers. Developments must plan and design their systems to expose all interfaces to developers in the outside world.

---

[445] Ilya Zaslavksy, "Readiness of Disciplinary Data Systems for Cross-Domain Interoperability within a Standards-Based EarthCube Reference Framework," National Science Foundation EarthCube Expression of Interest 1051, 2012.
[446] Tim Ahern, "A New Capability: Service Based Vertical and Horizontal Integration Platform," National Science Foundation EarthCube Expression of Interest 1041, 2012.

BUILDING A SOFTWARE ENGINEERING AND INFORMATICS COMMUNITY OF PRACTICE AND A SOFTWARE FOUNDATION FOR THE ADVANCEMENT OF EARTHCUBE

College of Charleston (1022): Proposes to build on the progress from the November 2011 Charrette to establish a Community of Practice (CoP) as a more formal collaborative software engineering and informatics community that is informed by the geoscience community, and that lays the foundation for a software governance organization for EarthCube.[447] This CoP would provide human and software resources and best practice guidance, including from domain scientists, for EarthCube projects from conceptualization and formation through completion. This CoP would also drive the creation of a governance organization for EarthCube to support sustainable software development that models other successful governance organizations, such as the Apache Software Foundation (ASF). This governance organization (tentatively named the EarthCube Software Foundation - ESF) would provide vehicles, such as an on-line presence featuring resources, people, interactive match-making for proposals, blogs, experience reports, and forums, for the CoP to engage members of the EarthCube science community whose proposals contain significant software development components.[448]

EARTHCUBE GOVERNANCE COMMUNITY

A larger virtual community with a stake or interest in governance was engaged in the roadmap writing process. Several means of communication were established so that community members' opinions, suggestions, and questions about the ideas coming forward were taken into account. The community feedback process was open to all, including those outside the EarthCube initiative. We recruited additional engagement through virtual webinars and through the use of the social media via the EarthCube Governance Ning Site.

WORKSHOPS

We conducted a series workshops between March 13 and June 1, 2012 of different working groups and stakeholders, covering a range of topics, as well as two plenary workshops, albeit offered at multiple dates and times, to allow more interaction and to accommodate participants schedules.

The E-Gov Steering Committee also met face-to-face, April 4-5, 2012, in Denver, Colorado to better establish rapport among the members who generally did not know each other, and to work out the details of the workshop goals, processes, and agendas. The EC-Gov Steering Committee (SC) developed a draft description of the goals of EarthCube governance during the April 4-5 planning meeting in Denver, drawing from the research and compilation carried out by the project's staff and Community Coordinators prior to the meeting. *These goals were distributed to the EC-Gov Forum and discussed virtually (online, webinars, or conference calls as needed) as the basis for the Plenary Workshops*. Planning and assigning of tasks for the Plenary Sessions were also an important components of April 4-5 meeting.

The E-Gov Steering Committee will also meet face-to-face on June 11, 2012 to finalize the Governance Roadmap and to engage liaisons with other EarthCube Working Groups and Concepts in anticipation of the June NSF Charrette.

PLENARY WORKSHOPS

[447] Bowring, Jim, "Building a Software Engineering and Informatics Community of Practice and a Software Foundation for the Advancement of EarthCube," National Science Foundation EarthCube Expression of Interest 1022, 2012.

[448] Sky Bristol, discussion post on EarthCube Governance Ning site, http://EarthCube.ning.com/group/governance/forum/topics/summary-of-governance-eois, February 29, 2012.

We held two Plenary Workshops online via webinar to introduce the communities to the goals, plans, and timetable for the EC-Gov project. The same workshop was repeated over a 7 day period to 1) Keep the attendance small enough in each event to allow and encourage effective participation by all those online, and 2) Provide opportunities for the widest possible participation. The agenda both Plenary Workshops contained the following governance-related topics: 1) Purpose and Scope of the Governance Roadmap Workshop; 2) Background and Research on Governance; 3) How We Jumpstart the Planning Process; 4) Governance Examples; 5) Governance Functions; 5) Broader Impacts & Linkages to Other Communities; 6) Governance Use Cases/Mental Exercises; and 7)Community Discussion. The Community Discussion portion of the Plenary Workshops generated much community feedback that was applied to subsequent events, and is a core component of the Governance Roadmap. Community feedback in its entirety can be found in Appendix 4.

TOPICAL AND DOMAIN WORKSHOPS ("VIRTUAL BREAKOUTS")

Given the complexity, scope, and diversity of issues that have to be addressed by this project, and in such a short time, virtual breakout sessions were conducted with the aim of gathering governance feedback and information regarding the 10 guidance points developed by NSF. Virtual breakouts were targeted towards specific groups within and without the EarthCube community, including EarthCube Working Groups and Concepts Teams liaisons, and industry, international, and government agency representatives. Reference materials disseminated in advance to participants via email and Ning discussion posts included basic information about EarthCube and governance, in addition to asking specific questions about governance:

1. Your past experiences with governance (Decision-making structures, procedures, communication, leadership, community-building, etc.): What's worked? What hasn't worked?
2. How do you define and envision EarthCube governance?
    1. What do you **need** in terms of EarthCube governance? What do you **want**?
    2. What are your **assumptions** about EarthCube governance? For example, are you assuming a consensus-based model, the formation of a technical committee for standards, or something else?
3. What you **NOT** WANT or DO **NOT** NEED in terms of governance?
4. What are some barriers to your participation in EarthCube? If any.
5. Remaining issues
    1. What is it that we're trying to govern? Who decides this?
    2. What issues does an EarthCube governance framework need to address

The virtual breakout sessions generated excellent community feedback, and often led to follow-up governance discussions via email and the EarthCube Governance Ning site.

Our original Expression of Interest indicated a desire to conduct breakout sessions with groups that were not initially targeted by the EarthCube community, for example, workforce education and training. Fortunately, most of the gaps initially targeted by the focus groups were filled by additional NSF led community groups.

We solicited written feedback from EarthCube Working Groups and Concept Teams to map their governance framework to Weill and Ross' Governance Matrix.[449] We asked each group to map the types of decisions that must be made (top row) to the governance archetypes (left column) (Table 25).

Decision **Archetype**	***IT principles***	***IT architecture***	***IT infrastructure strategies***	***Application needs***	***IT investment***
Domain monarchy					
IT monarchy					
Federal					
Duopoly					
Feudal					
Anarchy					
Meritocracy[450]					

TABLE 25: GOVERNANCE AS A SYSTEM[451]

Types of IT Governance Decisions:

- IT Principles
    - What is the enterprise's operating model?
    - What is the role of IT in the business?
    - What are ID-desirable behaviors?
    - How will IT be funded?
- IT architecture
    - What are the core business processes of the enterprise?
    - How are they related?
    - What information drives these core processes? How must the data be integrated?
    - What technical capabilities should be standardized enterprise-wide to support IT efficiencies and facilitate process standardization and integration?
    - What activities must be standardized enterprise-wide to support data integration?
    - What technology choices will guide the enterprise's approach to IT initiatives?
- IT infrastructure
    - What infrastructure services are most critical to achieving the enterprise's strategic objectives?
    - For each capability cluster, what infrastructure services should be implemented enterprise-wide and what are the service-level requirements of those services?
    - What is the plan for keeping underlying technologies up to date?
    - What infrastructure services should be outsourced?
- Business Application needs
    - What are the market and business process opportunities for new business applications?
    - How are experiments designed to assess whether they are successful?
    - How can business needs be addressed within architectural standards? When does a business need justify an exception to a standard?
    - Who will own the outcomes of each project and institute organizational changes to ensure the value?

[449] Peter Weill and Jeanne Ross, *IT Infrastructure: How Top Performers Manage IT Decision Rights for Superior Results*, Boston, MA, Harvard Business School Press, 2004, 11.
[450] Ross Gardler and Gabriel Hanganu, "Governance Models"
[451] Adapted from original governance matrix in Weill and Ross, *IT Governance*, 2004, 11.

- IT Investment and prioritization
  - What process changes or enhancements are strategically most important to the enterprise?
  - What are the distributions in the current and proposed IT portfolios? Are these portfolios consistent with the enterprise's strategic objectives?
  - What is the relative important of enterprise-wide versus business unit investments? Do actual investment practices reflect their relative importance?

We then expanded on Weill and Ross' original matrix to include a greater variety of decisions that must be made in a complex entity such as EarthCube. Types of decisions that we added include organizational principles and goals, organizational investment and prioritization, business model, and science application needs (Table 26). Although all the types of decisions that must be made in EarthCube may not be reflected in this matrix, it is a good start to the basic Governance Roadmap principle that EarthCube governance must work as a system of different models for different types of decisions that must be made, at different levels.

Governance Archetype/Model **Decision**	***(Business) Monarchy***	***IT Monarchy***	***Federal***	***Duopoly***	***Feudal***	***Anarchy***	***Meritocracy***	***Technocratic Utopianism***
IT Principles								
IT Architecture								
IT Infrastructure Strategies								
Business Application Needs								
IT Investment and Prioritization								
Organizational Principles and Goals								
Organizational Investment and Prioritization								
Business Model								
Science Application Needs								

TABLE 26: EXPANDED GOVERNANCE MATRIX[452]

Expanded Governance Functions:

- Science Application needs
  - What is the overarching scientific vision?
  - How do you meet scientific and community needs?
  - How do you determine if the organization is successful in meeting those needs?
  - Who will own the outcomes of each project and institute organizational changes to ensure the value?
- Organizational Principles and Goals
  - How do you determine the overarching principles and goals of the organization?
  - How do you measure progress towards meeting those goals?

[452] Adapted from original governance matrix in Weill and Ross, *IT Governance*, 2004, 11. Due to space constraints, governance archetypes were moved to the top row, while types of decisions were moved to the left column.

    - What processes allow to evaluate and organizational goals when necessary?
    - How do you ensure that the governance framework is aligned with the organizational goals?
    - How do you ensure integration of all the IT and scientific components, within social and cultural constraints, for an effective CI?
- Organizational Investments and Prioritization
    - What processes determine where resources are invested and if those investments are consistent with the enterprise's strategic objectives?
    - What changes or enhancements are strategically most important to the enterprise?
    - What is the relative important of CI-wide versus parochial investments? Do actual investment practices reflect their relative importance?
- Business Model
    - What processes determine the initial business model?
    - What processes provide for the evolution of the business model?
    - How do you determine which governance archetypes are applied to different components of the organization?

RESEARCH REVIEW

The Roadmap writing-process was informed by an extensive background research review on governance. The content of the research review was guided by input from Steering Committee meetings, virtual plenary sessions, and discussion posts on the EarthCube Governance Ning site. The review focused on governance definitions , concepts, models and mechanisms, historical infrastructure building, best practices and lessons learned, governance case studies, governance recommendations from informatics and government reports, investigations on community-building, and governance components of NSF EarthCube White Papers (Governance and Designs categories, submitted fall 2011). While the research review is in no way exhaustive, it provided a solid theoretical and concrete foundation that greatly informed the roadmap-writing process.

All notes, summaries, PowerPoint presentations, and reference lists were posted to the EarthCube Governance Ning site to foster community review and comment. Community members used the discussion forum to suggest additional resources, many of which were incorporated into the background research review.

DISSEMINATION OF WORKSHOP RESULTS

We employed multiple facets of information dissemination to distribute workshop results in the spirit of promoting transparency and engaging the broader community. The EarthCube Ning site provided ready access to an extended core community and served as the principal communication venue. It served as an online community repository for roadmap drafts, agendas, research reviews, and community discussion questions.

Interaction between participants was carried out through the implementation of various online collaboration environments including an EarthCube Governance Forum and an EarthCube Governance Steering Committee list-serv. List-serv membership was open to the public, and subscriber information was posted to the EarthCube Governance Ning site. Google docs was used for all collaborations regarding virtual breakout session topics and participation, and writing of the Governance Roadmap. In addition, webinars and virtual workshops were recorded and uploaded to an EarthCube Governance Vimeo site, which was linked back to the Ning site. In addition, we set up and EarthCubeGov Twitter hashtag and handle to easily track and monitor real-time conversations.

Mechanisms for actively promoting participation by all interested parties continues with the EarthCube Governance Steering Committee with extensive representation of geoscience and IT communities. In turn, each Steering Committee member has their own extended network of informal and formal partners and collaborators that will become the initial core of the EarthCube Governance Forum.

WORKSHOPS & MEETINGS HOSTED

All meetings were open to the public with agendas and Webinar call-in information posted to Ning:

- Governance Steering Committee Call, March 23, 2012
- Governance Steering Committee Call, March 30, 2012
- Governance Steering Committee Meeting, Denver, CO, April 4-5, 2012, 12 participants
- Governance Steering Committee Call, April 6, 2012
- Social Media Conference Call, April 9, 2012, 4 participants
- Virtual Governance Plenary Session, April 11, 2012
- Governance Steering Committee Call, April 13, 2012
- Virtual Governance Plenary Session, April 17, 2012
- Social Media Conference Call, April 19, 2012
- Workflow Working Group Weekly Call, Governance Briefing, April 19, 2012
- Governance Steering Committee Call, April 20, 2012
- Governance Steering Committee Call, April 27, 2012
- Governance Steering Committee Call, May 4, 2012
- Governance Steering Committee Call, May 11, 2012
- Governance Steering Committee Call, May 18, 2012
- Governance Steering Committee Call, "Roadmap Review," May 22, 2012,
- Governance Steering Committee Call, May 25, 2012
- Focus Group, "Engaging International Partners on Governance (Pacific Rim) Webinar," May 30, 2012, 5 participants
- Focus Group, "Engaging International Partners on Governance (Europe) Webinar," May 31, 2012, 6 participants
- Focus Group, "Engaging Industry Partners on Governance Webinar," May 31, 2012, 4 participants
- Focus Group, "Engaging Federal Agency Partners on Governance Webinar," May 31, 2012, 8 participants
- Governance Steering Committee Call, June 1, 2012
- Governance Steering Committee Call, June 8, 2012
- Governance Steering Committee Meeting, Roadmap review with EarthCube Working Group and Concept Team Liaisons, June 11, 2012
- EarthCube Charrette, June 12-14, 2012

COMMUNITY INPUT: EARTHCUBE WORKING GROUPS AND CONCEPT TEAMS

ALVA COUCH
TUFTS UNIVERSITY AND CUAHSI
X-DOMAIN INTEROPERABILITY CONCEPT TEAM REPRESENTATIVE
MAY 29, 2012

1. Your past experiences with governance (Decision-making structures, procedures, communication, leadership, community-building, etc.) What's worked? What hasn't worked?

I am most familiar with CUAHSI governance, including the CUAHSI Board, the Informatics Standing Committee, and the Users Committee. Decision structures are organized around general policies, software policies, and user priorities, respectively. This works well.

2. How do you define and envision EarthCube governance?

This question was left blank.

3. What do you need in terms of EarthCube governance? What do you want?

The x-domain interoperability group envisions two kinds of governance:

- scientific governance, which determines scientific needs and priorities.
- technical governance, which determines shared standards and technologies.

Scientific governance includes decisions as to what services will be made generally available, based upon *scientific* merit. Technical governance determines the most convenient technical ways to accomplish that.

3. What are your assumptions about EarthCube governance? For example, are you assuming a consensus-based model, the formation of a technical committee for standards, or something else?

I am assuming that the two governance groups above are representative committees of the community, perhaps elected by the community.

5. What you NOT WANT or DO NOT NEED in terms of governance?

I think it is important to separate technical issues from priorities, as described above. Scientists should set priorities, and tech leads should define and manage technologies and their lifecycles.

6. What are some barriers to your participation in EarthCube? If any.

The principal barrier to participation in EarthCube is that the domain groups (including CUAHSI) continue to have a primary allegiance to specific domains. I am concerned that EarthCube could compromise the primary missions of the domain-specific groups, by siphoning resources away from serving the needs of domain users.

Remaining issues

1. What is it that we're trying to govern? Who decides this?

In the case of X-domain interoperability, there is a natural workflow that starts with compelling use cases (strategic, special-purpose uses) that are generalized into reusable high-level services (general-purpose resources. The primary role of scientific governance is to choose which services should be made generally available, based upon the *scientific* evidence for or against the use cases.

Meanwhile, technical governance concerns how standards should be formed and how they should evolve. This is a matter of version control, lifecycle management, and technical decision-making.

2. What issues does an EarthCube governance framework need to address

Generalizing from specific success stories to develop new general-purpose services.

REAGAN MOORE
DATANET FEDERATION CONSORTIUM
LAYERED ARCHITECTURE CONCEPT TEAM REPRESENTATIVE
MAY 25, 2012

1. **Your past experiences with governance (Decision-making structures, procedures, communication, leadership, community-building, etc.) What's worked? What hasn't worked?**

   The DataNet Federation Consortium (DFC) is an NSF project tasked with implementing national data cyberinfrastructure. The approach is based on a "bottom-up" federation of existing data management infrastructure. This requires a focus on interoperability, the implementation of virtualization mechanisms that provide a uniform view across heterogeneous community resources, and the creation of a collaboration environment that enables researchers from multiple institutions to jointly participate in research initiatives. National scale projects participating in the DFC include Ocean Observatories Initiative, Hydrology research, iPlant Collaborative, Engineering design, Cognitive Science, and Social Science.

   The DFC focus areas have very similar requirements to the EarthCube initiative. Within the Layered Architecture Concept Award, the lessons learned from the DataNet Federation Consortium are being incorporated into the design of a framework for linking GeoScience community resources to collaboration environments. The goal is to build upon existing infrastructure, while enabling incorporation of new technologies, and enabling future research initiatives to build upon current research results. Since current research is conducted within a collaboration environment, a mechanism is needed to transition a collaboration environment into a community resource for use by future researchers.

2. **How do you define and envision EarthCube governance?**

   There are multiple national scale data cyberinfrastructure projects, each with a different set of challenges. For instance, the OOI manages real-time sensor data. The National Climatic Data Center manages climate data records. The two systems need to interoperate to enable deposition of OOI climate data records into NCDC. The two systems, however, are based on different infrastructure.

   The target for governance needs to be on interoperability. What strategies promote interoperability? What standards will enable improved interoperability? What extensibility mechanisms are needed to enable interoperability at the process level, the metadata level, the data level, and the user level? One way to characterize interoperability is through identification of name spaces for objects that are being shared, the identification of the operations that are performed upon each name space, and the virtualization mechanisms that are needed to ensure the operations can be performed across existing community resources. Within the DFC, the name spaces that are managed include users, data, collections, storage resources, state information (metadata), policies, and procedures. The document "Introduction to Layered Architecture" lists the associated operations and virtualization mechanisms.

There are at least two types of governance: Managerial and Technical. Managerial governance defines the policies for interactions between funding agencies and communities of practice that implement standards and technologies. Technical governance defines the policies and procedures that implement a community consensus on assertions about shared collections and processes. The DFC has chosen to automate Technical governance. Assertions about the community consensus are turned into properties that each object in a collection should possess. Policies are defined to ensure the desired properties are enforced. The policies are cast as computer actionable rules that control the execution of procedures. The procedures are created by chaining computer executable functions(micro-services) into workflows that are executed at remote storage locations. Each consensus is then characterized by a set of policies and procedures. Each user community may choose to enforce a different set of policies and procedures. There is no single set of Technical governance policies. However there is a policy-based framework that enables each community to apply their specific policies.

Instead of a data life cycle, the focus is on a community-based collection life cycle. A collection defines the context for research data. Within a research project, the context may be minimal, since the members of the research team have tacit knowledge on how the data were generated, what processes were applied, and the properties of the individual files such as naming convention. When the user community is broadened, say through a data grid that enables other research teams to collaborate on analysis of the data, the tacit knowledge must be made explicit. This requires transition from governance policies for a research collection to governance policies for a data grid (shared collection). The tacit semantics, data formats, and processing algorithms now need to be defined for use by the additional research team. When the collection is made public for use by the entire discipline, the user community broadens again. Now expectations for the collection context must be mapped to the community consensus on semantics, descriptive metadata, and data formats. When the collection is preserved in a reference collection, the user community broadens again to include future researchers. The context must now include the information and knowledge that defined the original research question, the governance that controlled the approach to solving the research question, and the metrics for success.

Within a community-based collection life cycle, the governance policies evolve to track the consensus of each broader user community. This is a direct measure of the NSF mandate to track broadened impact of research results.

There are multiple user communities: Researchers in national scale projects, Researchers in the "long tail" of science, and Students. They all face the same challenges with quantifying context for research data and research results. There is generic infrastructure that enables each community to provide the needed context. One can build either a Student digital library, or a project data management system, or a collaboration environment with a policy-based data management system. The policies and procedures may be different in each environment, but the underlying data, information, and knowledge virtualization mechanisms are the same.

3. **What do you need in terms of EarthCube governance? What do you want?**

For Technical governance we need mechanisms to define a community consensus on desired properties, policies, and procedures.

For Managerial governance we need mechanisms to ensure sustainability of the infrastructure. This includes evolution of the infrastructure to incorporate new technologies, new approaches towards quantification of knowledge, and new approaches for governance.

4. **What are your assumptions about EarthCube governance? For example, are you assuming a consensus-based model, the formation of a technical committee for standards, or something else?**

   I strongly urge a consensus-based model. However the consensus is relative to the community that is exploring a research topic. There are multiple levels of governance. The entire discipline may select the next 10-year research challenges. National scale projects may implement the required infrastructure to collect data, implement solution methods, and make the capabilities available to the discipline. Researchers within the "long-tail" may compare local information with the national initiative results and refine, explore, develop extended result sets. Students may participate in research initiatives under appropriate policy controls. The governance for each of these communities is different.

   Standards may promote interoperability for a specific operation on a specific name space. However, new operations continue to appear, new data types are created, new features within data sets are recognized and become relevant, etc. Standards focus on running code, not on new ideas and new research challenges. The governance model needs to allow new standards and new interoperability mechanisms within the infrastructure.

5. **What you NOT WANT or DO NOT NEED in terms of governance?**

   Governance metrics that are not quantifiable are not needed. For Technical governance, it should be possible to verify each desired property of a collection. For each governance policy, there should be an associated validation policy.

6. **What are some barriers to your participation in EarthCube? If any.**

   Instead of barriers, there are opportunities to promote interoperability mechanisms.

   Remaining issues:

7. **What is it that we're trying to govern? Who decides this?**

   Governance should ensure reproducible science. At a minimum, the information and knowledge required to reproduce a research result has to be provided for each research project. Governance is the process to make this feasible through provision of the necessary infrastructure. If a project is funded, infrastructure must either be provided or developed to enable others to check the result.

   Each research community makes assertions about their research results. A research community should be able to verify each assertion. Governance needs to ensure each community has the infrastructure to validate their own assertions.

8. **What issues does an EarthCube governance framework need to address**

   The dominant issues are the provision of mechanisms for groups to interoperate (reach consensus), and mechanisms for technology to interoperate (virtualization mechanisms).

Additional Governance Input

From the perspective of the iRODS data grid, governance is accomplished through the following steps. Each collaborating group builds a consensus for:

- purpose of collaboration (shared collection, shared research initiative)
- properties that the collaboration environment will maintain (completeness, integrity, standards, access)
- policies that enforce the desired properties (computer actionable rules)
- procedures that derive the desired properties (computer executable workflows)

- state information that tracks results of applying procedures
- assessment criteria that verify desired properties have been enforced.

We implement this approach using a distributed rule engine, a rule language, and workflows composed by chaining basic functions. At the moment, the scale is:

- 74 policy enforcement points (locations in the middleware where we check whether a policy needs to be enforced)
- 250 micro-services that encapsulate basic operations
- 209 system state variables to track outcomes (plus arbitrary number of user-defined status flags, descriptive metadata, provenance metadata)

REAGAN MOORE, DATANET FEDERATION CONSORTIUM, LAYERED ARCHITECTURE CONCEPT TEAM REPRESENTATIVE MAY 31, 2012

These comments are based on the DataNet Federation Consortium, an NSF funded project to build national data cyberinfrastructure for NSF funded research projects. Current participants include Ocean Observatories Initiative, Hydrology projects, and the iPlant Collaborative. We expect policies to be developed independently by each project for governance, management, administration, and assessment. The data grid (iRODS) allows each group to enforce their own policies, independently of the other projects. The infrastructure is designed to be highly extensible, separating policy enforcement from data manipulation, data storage, and data access mechanisms. While the software infrastructure is generic, the policies and procedures can be unique to each user community.

For collaboration environments, the policies need to represent a consensus across the participating groups.

Decision	*IT principles*	*IT architecture*	*IT infrastructure strategies*	*Application needs*	*IT investment*
Archetype					
Domain monarchy					
IT monarchy					
Federal			*X*		*X*
Duopoly					
Feudal	*X*	*X*		*X*	
Anarchy					

TABLE 27: DATANET FEDERATION CONSORTIUM GOVERNANCE MATRIX

***IT Governance* details three questions that effective IT governance must address:**

1. **What decisions must be made to ensure effective management and use of IT?**
   - We think in terms of the driving purpose for the shared collection.
   - The purpose is translated into a set of properties that should be enforced over time.
   - The properties are enforced through management policies, which are cast as computer executable rules.
   - The rules control the execution of procedures, which generate the desired properties.
   - The procedures create state information which is automatically saved.

- The state information can be queried by periodic policies to verify that the desired properties have been maintained.

This automates management and administrative tasks, and provides a way to validate the system. We strongly believe that automation of policy enforcement is an essential component of large-scale data management systems.

2**. Who should make these decisions?**
The community that is promoting the shared environment should make the decisions. The policies can be augmented by requirements from funding organizations (data management plans), legal requirements (HIPAA, FERPA), and institutional commitments (resources).

3. How will these decisions be made and monitored?
The consensus is the difficult step. Fortunately, policy-based data management systems support the evolution of the policy data sets. If the initial policy is insufficient, it can be updated without having to modify code. Versions of policies can be applied, and collections can be migrated from an original set of policies to a new set of policies.

<u>IT defines what decisions should be made:</u>
1. IT principles - Clarifying business role of IT
The business objective can be cast as a set of metrics that are evaluated over time for use, customers, payment. Given a set of metrics, policies and procedures can be defined to evaluate the metrics. The persons who make the decisions will evolve over time.

2. IT architecture - Defining integration and standardization
We observe that a highly extensible infrastructure is needed, since new types of data, new types of services, new types of storage systems, and new technologies will emerge during the lifetime of a project. For policy based data management systems, the extensibility is ensured by decoupling client access mechanisms from storage. Clients interact with middleware, which on behalf of the client interacts with the remote resource. The types of access mechanisms range from web services, to web browsers, to Unix tools, to Grid tools, to portals, to digital libraries, to workflow systems, to I/O libraries, to load libraries, to file system interfaces, to synchronization tools. The same challenge occurs with storage systems, which may use I/O protocols from Windows file systems, Unix file systems, tape archives, web sites, FTP, object-based systems, databases. The middleware translates from the client protocol to the storage protocol.

3.IT infrastructure - Determining shared and enabling services
Control of sharing is supported through policies that evaluate constraints between any two name spaces (files, users, storage systems, state information). Pinning of data to a storage device is a constraint imposed between users and storage. Browsing control is a constraint between users and state information.

In a policy-based data management system, each client action is trapped at a policy enforcement point. A rule base is then checked for policies that control the action, and executed by a distributed rule engine. This makes it possible to control every operation. Services are enabled by asserting an appropriate policy.

4. Business application needs - Specifying business need for purchased or internally developed IT applications

If generic policy-based data management infrastructure is used, business application needs can be separated into the set of policies and procedures required by the business application. The intellectual property then resides in the set of policies, not in the software infrastructure. Since the policies can be applied uniquely to a given collaboration, business needs can be directly addressed, controlled, and implemented.

5. IT Investment and prioritization - Choosing which initiatives to fund and how much to spend

The differentiation between generic infrastructure and project specific policies and procedures enables investment at multiple levels. A federal agency can support the development of the generic infrastructure. A commercial enterprise can support the implementation of business specific policies.

The development cost for generic infrastructure is non-trivial. The SRB storage resource broker cost about $15 million to develop, apply, maintain, and support over 10 years. The iRODS data grid is slightly cheaper, costing about $6 million over 6 years.

And who should make those decisions:

1. Business monarchy: top managers – the decision can be whether to use generic infrastructure, or role your own
2. IT monarchy: IT specialists – can decide which resources will be integrated (storage systems, compute servers)
3. Feudal: each unit makes independent decisions – can pick the policies and procedures
4. Federal: combination of corporate center and business units with or without IT people involved – can define the investment model
5. IT duopoly: IT group and one other group (ex: top management or business unit leaders) - can pick specific policies and procedures

Anarchy: isolated individual or small group decision making – can pick specific policies and procedures

---

SHAHANI MARKUS WEERAWARANA, PH.D.
COMPUTER SCIENTIST / VISITING LECTURER / VISITING RESEARCHER, INDIANA UNIVERSITY
WORKFLOW COMMUNITY GROUP MEMBER & INTERNATIONAL COMMUNITY REPRESENTATIVE

I am involved in the Workflows Community Group as a visiting researcher at Indiana University. I am very interested governance issues and decided to follow your Webinar today (since I am originally from the Asia/Pacific region...).

Although I posted a comment during the Webinar, it was not possible to share more thoughts and comments - hence I decided to send them to your mailing list. Many of these issues may have been extensively discussed within your community group and I apologize if I am rehashing old stuff.

My background in governance is primarily from the eGovernment domain as well as from the governance in SOA (service-oriented architecture) area and through the management of a large IT organization.

- As you identified in your presentation, governance spans many areas, (as well as levels). Yet, it appears that each governance area requires specialized mechanisms/infrastructure/policies etc. For example,
    - Information/data governance would require attention on many aspects including,
        - information ownership model along with associated ownership roles & responsibilities
        - definition and identification of information security levels and associated access & usage policies
        - information provenance guidelines and enabling infrastructure mechanisms
    - IT/Cyberinfrastructure governance would require attention on an almost completely different set of aspects (which were quite comprehensively listed in your presentation).
    - *Thus your roadmap may need to include the process of identifying governance areas relevant to EarthCube and the specific governance aspects of each area.*
- Every aspect of governance should include a temporal component inclusive of associated life-cycles (the process of originating, maturing and retiring particular governance components)
- Careful attention must be given to the legal aspects of governance.

- All communication/collaboration plans that are specified in the EarthCube Governance Framework should include explicitly defined "escalation paths".
- If EarthCube will be expanding to include significant global collaboration, all governance aspects should consider cultural impacts. For instance, if a particular governance area were to adopt the meritocracy governance archetype, due consideration should be given to culturally induced "silent/unseen work ethics".
- There should be clear differentiation of strategic vs. operational governance structures. The operational structures would necessarily have a greater monitoring & management flavor, whereas the strategic structures would be more focused on guidance and decision-making.

SHAHANI MARKUS WEERAWARANA, PH.D.
COMPUTER SCIENTIST / VISITING LECTURER / VISITING RESEARCHER, INDIANA UNIVERSITY
WORKFLOW COMMUNITY GROUP MEMBER & INTERNATIONAL COMMUNITY REPRESENTATIVE

Elaboration on previous email:

Every aspect of governance should include a temporal component inclusive of associated life-cycles (the process of originating, maturing and retiring particular governance components) **Do you have any specific governance components in mind?**

Consider governance of policies and procedures. There should be a clearly defined timeline for the definition, operation and retirement of such artifacts. Yet beyond them, practically any component being governed has an associated life-cycle. As an example, consider a project team or a committee. Similar to the Apache Software Foundation (http://www.apache.org/), there is a life-cycle process of incubation, graduation and retirement. By associating a pre-specified timeline with such life-cycles, you can ensure timely reviews to ascertain if a component is ready to move on to another stage or not. In the case of named/numbered artifacts, it may result in version changes.

Thus in general, when putting together a governance framework, it would be useful to require "thinking about" the life-cycle of each aspect/component/artifact and the associated timeline. This will help ensure timeliness, relevance and "liveness" of a governance framework.

Careful attention must be given to the legal aspects of governance. **What legal aspects are you thinking of?**

There are legal overtones to many governance activities. Consider something as simple as an EarthCube participant collaborating via a mailing list. What are the IP rights that are applicable on each communication? What are the legal strictures that apply to participants in community groups and project teams? What are the legalities present in policies and procedures? How can the governance framework ensure that international trade laws or global IP rights are not violated in the process of collaboration among global EarthCube members?

So a governance roadmap should highlight the importance of providing legal guidelines. Most EarthCube participants would be completely unaware of legal aspects and it is up to the Governance framework to protect them, their community, and EarthCube as an entity.

All communication/collaboration plans that are specified in the EarthCube Governance Framework should include explicitly defined "escalation paths". **What do you mean by 'escalation paths'?**

Many communication and collaboration plans are very detailed regarding regular activities that should happen. However, many forget to specify the escalation path when *things go wrong!* How can any member of the community create an "alert" to enable timely attention on a brewing risk? Who/What is responsible for following up on each alert. What is the escalation path that *leads out of or beyond the immediate communication/collaboration group boundaries*?

This would also be part of the overall risk mitigation plan in a governance framework.

ADDITIONAL INPUT FROM THE EARTHCUBE COMMUNITY ON SCIENTIFIC VERSUS TECHNICAL GOVERNANCE

"What has come out rather clearly in my study of interoperability is that there are two kinds of governance:

1. scientific governance, where the goal is to set priorities.
2. technical governance, where the goal is to set strategies.

I don't know if the dividing line between these two is so clearly drawn in other working groups, but in this group, there seems to be a difference.

So, my question is, should there be a distinction?"

"This distinction [scientific vs. technical governance] is important and should be part of EC governance. It is particularly critical since much of EC technology development will be distributed and will involve federated activities. However, I do think that there needs to be a single overarching entity like the Board of Directors or some other committee that will steer the entire EC effort.

Our experience in Unidata is consistent with that. We have a Policy Committee for overall policy decisions and governance, a User's Committee for community engagement and working with our scientific users and educators, and technical advisory committees for our software projects."

"[Scientific vs. Technical governance is] a useful separation of tasks. For example, attention to SOA, software engineering practices, and data exchange standards would fall more under technical governance, while deciding what problems to solve would fall under scientific governance."

"The DDMA group has also identified governance (but in the DDMA context) as an activity unto itself. As you can imagine, there are many governance issues at the data level that we feel we need to catch right at the functional level some, if not all, of which might bubble up to the larger governance structure for EC. Even at the DDMA level, I can see science vs. technology governance issues (e.g. what level of products should be generated vs. how does one decide where a given data technology is on the hype cycle, etc.)."

CARROLL HOOD
RAYTHEON

I was reading the Governance white papers and came up with the following observation: most of them advocated some form of a PMO function (often used different names, but the concept was consistent,) but a few advocated governance arising organically through collaboration. Some of the organizations doing so (RENCI, OGC, to name a couple) based their recommendations on what they use operationally; however, in my mind there is a distinction between collective organizations who simply collaborate/advocate and those who are involved in infrastructure development.

…We had a few simple recommendations:

1) EarthCube needs a proactive PMO (fostering integrated system development as well as cross-domain collaboration)
2) Any solicitation for a PMO ought to be constructed such that private sector organizations have a fair opportunity to prime (e.g., do not limit to academic or academic consortium-led groups only)
3) The PMO ought to be in place prior to the results of the initial prototypes/risk reduction activities being delivered.

We used the example of the GENI Project as an example of a successful private-sector led PMO

The fact of the matter is that private sector companies such as Raytheon will hesitate to participate in EarthCube unless there is some opportunities for us to contribute in a meaningful way. Unfortunately, in my world that includes the opportunity for bookings and sales. I know there are some in the academic world that view that perception as "robbing from the opportunity to do good science." We view it as providing the engineering rigor to ensure that the underlying infrastructure is not the constraint to performing good science. There are many, many instances in NSF lore where the FDR was deemed "successful" but the design was really not mature enough to initiate construction (MFREC). As a cross-discipline, highly distributed concept, EarthCube without a proactive and integration savvy PMO would like suffer that same fate.

We're not trying to take over, just trying to ensure that the playing field in level enough for us to play (and to encourage our bean counters to let us invest and contribute).

The PMO issue is only a fraction of the totality of Governance, but the solution to this issue (PMO or no PMO) will impact the design approach from the onset.

Does this make any sense?

I am available to discuss in more detail at your convenience.

Carroll Hood

Industry Input: Project GENI
"Attached is a copy of the original solicitation for the GENI Project Office that NSF released back in 2006. I am not suggesting that it is a verbatim blueprint for an EarthCube PMO solicitation, but it does offer some interesting points:

- the GENI PO provides oversight for the design, development and integration of the various piece parts
- pass-through $$ are included to enable the GENI PO to solicit mini-proposals to address key functional gaps or to mitigate risks in order to move the enterprise forward

This is the kind of proactive PMO that I believe is needed for EarthCube; however, I understand the need for the stakeholder community to arrive at a consensus on the appropriate governance model. This approach will only work if a top-down model is selected as optimal.

By all accounts, GENI is viewed as successful. Other NSF activities without a proactive PMO having integration responsibilities (e.g., NEON, OOI) have faced (or currently facing) construction-phase issues.

Note: this solicitation was won by BBN. I believe that there were at least seven competitors (at least 3 were private-sector led teams, and the rest were academic-led teams.) Note that the solicitation explicitly invited the private sector to bid.

As an aside, Raytheon acquired BBN in 2009. So we have amassed some experience and lessons learned in successfully managing an academic-based research initiative. This is not relevant to my overall objective (e.g., work to ensure that the private sector has a fair and viable opportunity to contribute meaningfully to EarthCube development), but in the spirit of full disclosure, I just wanted to put that out there."

(See attached file: GENI GPO Solicitation.doc)

COMMUNITY INPUT:
EARTHCUBE GOVERNANCE VIRTUAL PLENARY SESSION
APRIL 11, 2012

Community Discussion Questions and Comments

- Focus Groups
    - Question: how to get focus group started? (Jim Bowring)
        - Logistics – where to post information, how to organize, etc.
        - Want to focus on software
    - Response: (Lee Allison)
        - Post ideas to Ning and ask for community input
        - Find people who want to be involved
        - AZGS will provide staff support
            - Focus Group guidance document
                - What to address
                - NSF 10 guidance questions
                - Summarize
                - Time Table
            - Webinar logistical support
- EarthCube Vision
    - Dave Fulker
        - Question as to who determines overall goals needs further discussion – vision not necessarily dominated only by domain scientists
            - It is not helpful to set aside group as dominating the development vision
            - The Vision should be a combination of science and knowing what technologies are available to implement the vision
            - Scientists as researchers and educators
            - "The art of the possible"
        - Leadership
            - Is more concerned about group thinking yielding "impossibly jumbled" criteria than making sure every single person is include in decision-making process
                - There needs to be a way to define focus
                - The organization should be strong enough to say no
        - Will post documents on leadership that defines a clear focus
    - Kerstin Lehnert
        - Concerned about involvement of scientific community
        - Should have focus group on the involvement of the scientific community
        - How to do better outreach to the community
    - Mohan Ramamurthy
        - Concerned scientists won't be engaged enough
- Education and outreach should be important part of EarthCube
    - Interest in creating focus group
    - People Interested
        - Dewayne Branch, Chris Keen, David Arctur, Jim Bowring

COMMUNITY INPUT:
EARTHCUBE GOVERNANCE VIRTUAL PLENARY SESSION
APRIL 17, 2012

Community Discussion Comments and Questions:

Andrew Maffei

- Slides are complete. Should be review once we identify the roadmap.
- What do we mean by 'soft' ware, e.g. do vocabularies and ontologies fall into this or are they separate?
    - IETF is another governance group in IT Group
- When doing research it would be interesting to determine the number of paid staff the program has and what level of reliance there is on volunteer labor

Bill Michener: DataONE

- Three things missing / underrepresented
    - Governance scales by two things
        - Money
        - Size of organization and size of partners
    - Evolution of organization
        - Most start fairly simply
            - Use cases as a means to define scope and simplify the process
    - Roll of communication

Ilya Zaslavsky

- Propose focus group on software that supports governance requirements for software tools

Sayeed Choudhury

- Governance issues at various levels – EarthCube
    - Tiers of governance based on what level
- Good to review NSF report on virtual organizations

Bob Chen

- Governance issue associated with large and small science
- How do we integrate with those that already have governance structures?
    - Concern about bringing in larger organizations, like NASA

Cecelia DeLuca

- Model interoperability and data interoperability are two different things
    - Huge challenges if it's a big system
    - Good for a use case
    - Index for groups trying to promote interoperability - EarthCube needs something like this
    - Design-infrastructure
    - Catalogs, trackers, etc.
    - Schema that needs to be exposed for interoperability
        - Expose information on different projects

Kelly Rose

- DOE project leveraging EDx
    - Energy Data Exchange, NETL

Engaging International Partners on Governance (Pacific Rim); May 30, 2012

- Feedback: no feedback during the call
- Comment: Governance is necessary in many areas, but it is also necessary at different levels, that is, at the community level, project level, and team level, etc.

Engaging International Partners on Governance (Europe)," May 31, 2012

- Question: How will coordination occur with international participation?
- Answer: Not sure yet, but Belmont Forum for multi-national collaborations is a possibility; potential for EarthCube to be eligible for funding
- Comment: Scientific and technical governance issues related to some of the issues mentioned in the presentation (remaining questions and challenges, etc.)

Engaging Industry Partners on Governance, May 31, 2012

- Feedback: glad to have the "what are we trying to govern" chart - this also hits that we are not just governing IT, it needs to be holistic, point is to get grand challenge type science
- Feedback: if you fundamentally look at the grand challenges research, whatever people do to use EC as an enabler, it could stand the gambit, research questions could be validated or be led to the next science, the CI cannot be the failure, CI needs to be built with an engineering framework
- Feedback: existing NSF activities are having problems because the governance that works has an integration function; a lesson learned from other NSF activities is that somebody has to serve as a systems integration role to make the parts fit together so that the CI becomes the enabler so that it's a means to an end rather than an end itself
- Example: NSF project GENI, network office in place with a project office, funded it with $12m per year $2m for the governance structure, the other $10 is on proposals to mitigate risks to help the entire enterprise move forward
- Comment: our participation hinges on the assumption that there is an opportunity for the private sector to have a level playing field, if you don't have the engineering that enables science, then you don't have a complete system.

Engaging Government Agency Partners on Governance, May 31, 2012

- Question: Expand on the initiative?
- Question: What types of data are being governed?
- Comment: Suggestion for additional case studies - find the strategies that have lasted the longest and with the most results, i.e. the strategies that have made data the easiest to access, use, and maintain.
- Comment: Suggestion for reviewing the funding mechanisms of various organizations and study what's worked, what hasn't worked.
- Comment: Suggestion to include NSGIC - including the 50 GIOs.
- Potential for an Appendix to the Roadmap?
    - Documenting the webinars and consultations we've had with the other workgroups - who all was contacted

COMMENT BY SKY BRISTOL ON FEBRUARY 29, 2012 AT 8:03AM

There is an interesting mix of topics lumped together here under "Governance." I'm not sure on the reasoning behind this, and I unfortunately missed the webex presentation the other day where that might have been discussed. It appears from what I have heard and what's in the slides, that the intent is to see these four groups with the splashy icons come together on a combined proposal for a workshop and related activities to the tune of $100K each. We've been kicking this around a little bit so far in the DDMA group.

Here in this group, we seem to have a few fairly different ideas brewing. There are a couple of calls for EarthCube to either form an organization similar to the Open Geospatial Consortium or else simply work directly within the OGC as a governance mechanism for advancing standards and methods of importance to accomplishing the EarthCube mission. There are a couple of calls for a slightly different, though potentially compatible, community of practice model.

Then there are three EOIs that have to do with building actual cyberinfrastructure or tools in somewhat focused areas. I'm not completely sure why those ended up in this group, and I would love to hear the thinking behind that from NSF folks. The "Strategic Organizational Framework (Centralized Model)" from Hannes E. Leetaru and colleagues seems like it might be a much better fit within the DDMA group as it is proposing a centralized data archive for geoscience data.

The proposal for "Designing a community-based framework for modeling the consequences of climate change on coastal systems" from Gary Crane and colleagues does seem like a good fit here and might be an excellent candidate for a prototypical experiment of a governance model similar to what is proposed in the two related to the OGC model.

The summary written for "OpenQuake Social Networking for EarthCube from Geoffrey Fox and colleagues pulls out some key terms and phrases from the original EOI, but I question its placement in the Governance bucket. Sure, I think that EarthCube and the geoscience community at large could benefit from some focus on how collaborative information management tools in the research space really need to work, but I don't think that was the intent of this EOI. I would see it more in the area of data integration and semantics. That EOI seems to be getting much more at how we mine, integrate, analyze, and use for research and hazards response the massive amount of only partially structured information in the "public data cloud." That's a very different question, I think, than what most of the other EOIs in this space deal with.

The one interesting dovetail I see could come about through the whole exploration of collaborative information management. What if we could somehow create a federation across the many information platforms that we use in conducting research where we capture scientific workflows, data analysis and visualization methods, and other byproducts of our work? What if those byproducts be systematized in some way to enrich the education and research experiences across our diverse fields of study? There are applications like Kepler and VisTrails and aspirational things like Academic Room. But what if we started finding the common threads across the information that those types of things might advertise about the work going on within their confines? What if we took inspiration from what a whole bunch of people have figured out how to do in federating simple information across social networks and apply that to the scientific domain, defining our own common threads that crystalize the just in time information a researcher needs to have when exploring a particular problem?

Again, I think there are some folks over in the Semantics and DDMA areas that are dealing with this stuff. If you're not familiar with it, you ought to check out what folks in the Semantic Web Cluster in ESIP are up to in terms of simple linked data standards and implications to solving the still elusive citation metadata problem.

I am interested to hear from NSF folks who put this "cluster" together on the reasoning behind some of the inclusions and from folks put together into this group on how you all think these fit and where we go from here. P.S. Sorry for the rambling; it's early, my morning espresso has kicked in, and I'm finally feeling human again after being sick for several days.

Reply by Clifford A Jacobs, moderator on February 29, 2012 at 12:26pm

Hi Sky,

Thank you and your colleagues for participating in this discussion group. I'd like to respond to your questions about how NSF chose the groups and their participants. I think we agree that the EoI's provided a window into the complexity of the geosciences milieu, and that the boundaries within the complex environment are not well defined and subjective. In part, our goal in setting up the community groups was to be as broadly inclusive as possible and to provide a context to initiate a comprehensive dialog around what appear to be four broad themes. Your point is well taken that some of the EoI's that were grouped together here address the governance issue directly, while others address the subject peripherally and might be better suited in another discussion group or groups. If there is general consensus on that point, we are happy to work with you to facilitate that. From the outset, we have encouraged the community groups (workshop proposals) and the concept awards (EAGER proposals) to initiate and sustain a collaborative dialog. Group membership was initially selected by NSF, and is meant to be flexible if members believe they can better contribute to the activities of another group. We continue to encourage the EarthCube community at large to participate in any or all the activities NSF will fund.
NSF turns to the community for their guidance in identifying the complexities that must be confronted in realizing the EarthCube vision and detecting the assets (technical, human, social) that are available, or not, that can be employed in meeting the EarthCube challenge.

Reply by Sky Bristol on February 29, 2012 at 12:37pm

Hi Cliff,

Thank you for the response. You filled in a valuable missing link for me that I had heretofore missed somehow - the fact that you are thinking about both workshop proposals and concept awards. The diagram you had in the slides from the other day make a lot more sense to me now. It would be really cool for each of these thematic groups to help foster and bring along some prototypical R&D work that will introduce new capabilities or help move some key concepts along as a companion to the overall community building effort. That provides me with some better perspective in the grouping here and in the other areas. It also seems from the diagram that there isn't necessarily an inherent coupling between the concept type activities, leaving the possibility open for some super creative "fringe" group with great ideas to poke at the hornets' nest of the "establishment" and offer a useful perturbation in the system.

I really appreciate your comments, and I look forward to continued better understanding of the milieu.

USGIN STRATEGIC PLAN 2011

Posted by Stephen M Richard on February 28, 2012 at 4:42pm in Governance
Back to Governance Discussions
Result of advisory committee discussions on 5-year strategic plan for the US Geoscience Information Network.

NATIONAL GEOINFORMATICS COMMUNITY REPORT

Posted by Stephen M Richard on February 28, 2012 at 4:31pm in Governance

There have been a number of community workshops over the last several years discussing geoinformatics community building and governance. This is the first of several we'll be posting. Workshop on Working towards a National Geoinformatics Community (NGC) USGS Denver Federal Center, Denver, Colorado, September 23-24, 2010

Reply by Clifford A Jacobs, moderator on February 28, 2012 at 5:42pm

Hi Steve,

Thanks for identifying this report as one that has the potential for laying foundational work for EarthCube. As you know, the challenge is to expand the thoughtful and collaborative process that produce this, and other reports, to encompass all of geosciences. Although there are similar challenges throughout the geosciences as those identified in the NGC report, there are unique challenges particular to the various communities within the geosciences and often alternative solution approaches to common and unique problems. We look forward to the "listening" to the community dialog and facilitating discussion.

Cliff

A TECHNOLOGY-ENABLED MODE OF ORGANIZATION

Posted by Cecelia DeLuca on February 29, 2012 at 5:37pm in Governance

Hi all,

This may be the wrong place and time, but here's a somewhat different perspective.

Attached is a set of slides presented at the recent Ocean Sciences AGU that describes an NSF-funded project our group has been collaborating on, focused on governance in "communities of communities". We have gotten an incredible response from a variety of different groups, saying "yes, that's just what we need!". The idea is to create a space for either housing or "indexing" projects. Indexing means that we are standardizing and exposing the project metadata that is essential for collaboration and governance, such as the location of trackers and repositories, and processes such as how prioritization happens. We do this in a web environment that also offers access to a set of all-purpose community resources - an Earth System Grid data node, visualization services, etc. Key artifacts can be searched and saved back to the space as part of the metadata associated with the project. Very important, this environment also enables projects to be linked as peers and hierarchies, so that you will be able to get consolidated views (e.g. show me the trackers of my child and peer projects as well as my own). This is very much technology driving the possibilities for large-scale organization: we see disciplinary "clusters" starting to develop in the environment, which may also be cross-linked. We'd be interested in getting feedback on the slides and ideas, if they are of interest to EarthCube. Of course, the environment is also open to others for co-development (it's written in python and quite extensible). Please be warned, the project is still something of a baby, is a little rough around the edges, and is not at its permanent URL yet.

The big thing a structure like this does is link accessibility to information directly to governance. For large collaborations that must produce products, that linkage is central to the way governance needs to be structured - as argued in the slides, access to critical information is a prerequisite for collective and transparent decision making, which leads to trust ... which leads to functional organizations.

Best,
Cecelia

GOVERNANCE DISCUSSION OUTLINE - BASED ON A DECADE OF COMMUNITY BUILDING

Posted by Lee Allison on February 29, 2012 at 11:39am in Governance

Issues regarding governance in the geoscience cyberinfrastructure community have been discussed intensely for the past decade. I took an outline from a 2002 NSF-sponsored workshop on this subject and added a section on Workshop Processes as a suggested format for a new workshop proposal to NSF.

It seems that the decade of preparation has put us in a position to move forward now. The outline is attached.

Reply by Cecelia DeLuca on March 1, 2012 at 11:32am

Hi,

This looks like a sensible outline, but I would like to add two items that I think are major challenges. I am assuming that EarthCube will try to find out about and leverage the many elements of community infrastructure that already exist (regridding packages, data distribution facilities, data archival packages, data and metadata standards, etc.). Most long-standing projects have their own governance already and their own development timelines and sponsor requirements. As everybody knows, the challenge in putting together large scale infrastructure is combining such contributions. Executive level governance is important, but the finished product is made or broken on the quality and integration of components and on information availability and flow in a network of national and international partners. In a large software integration project, especially one that is distributed and lacks line management, organizing information (finding out where the trackers are - when the calls are happening - how priorities coming from multiple places will be balanced - what exactly will be in the releases of specific components) is one of the biggest challenges. The other huge challenge is creating norms for governance at the system component level. It can be incredibly difficult for a high level governance body to influence a critical system component whose local governance is non-existent or dysfunctional.

This view changes the emphasis in governance development somewhat from focusing on the strategic, Executive level aspects, to 1) paying close attention to how relevant information about constituent projects is found, exposed, and consolidated for practical use and 2) figuring out how to create governance norms at the constituent project level, and within the broader community. So I would add information flow and the creation of governance norms to the list of major challenges.

- Cecelia

INITIAL EARTHCUBE GOVERNANCE REFERENCES AND WORKS CITED LIST

Posted by Genevieve Pearthree on March 28, 2012 at 8:23pm in Governance

Reply by David Arctur on March 30, 2012 at 1:48pm
- Cargill, C. (1989) Information Technology Standardization: Theory, Process & Organizations. Digital Press, 252 pgs, ISBN 0-13-464470-0.
102 - Johnson, S. (2002) Emergence: The connected lives of ants, brains, cities, and software. Scribner, 288 pg. ISBN-13: 978-0-684-86876-9.

Earth Cube as a Double Loop Organization

Posted by Bruce Caron on April 4, 2012 at 5:08pm in Governance
While listening to the discussion this morning, it occurred to me that people might be talking about two parts of a larger governance model.
I've put some ideas into my blog about this: Earth Cube as a Double Loop Organization

Proposal for a Software Engineering governance road-map Focus Group
Posted by James F. Bowring on April 13, 2012 at 10:12am in Governance
We propose the formation of a software engineering governance road-map focus group. This focus group will produce a one page road map for directing the development of software governance for EarthCube. This road-map might include proposals from our initial EOI in which we identified two main governance areas for study. The first is the formation of a software engineering and informatics Community of Practice (CoP) for EarthCube. Such a collaborative CoP could provide a forum for software engineers, computational scientists, and domain scientists to exchange ideas, technology, and techniques directly applicable to the building of software infrastructure for EarthCube. The second governance area for study is investigating how to support sustainable software development for EarthCube. For example, we could consider how to form and evolve an entity such as the EarthCube Software Foundation (ESF), which might be the first product of the CoP. The ESF could be modeled after other successful software governance organizations, such as the Apache Software Foundation (ASF), which provides the infrastructure for the development and maintenance of software. The ESF would likely provide services such as: 1) Lifecycle support for the incubation and maturation of software projects for the EarthCube community; and 2) Sponsorship of conventions and best practices for software development, interoperability, and community adoption.

Reply by Chaitan Baru on April 13, 2012 at 2:49pm
I second this notion.
This is also consistent with an EarthCube white paper that we had submitted entitled, EarthCube TIVO: Test, Integration, and Verification Organization. We are unable to imagine how a system--even one as widely distributed and loosely federated as EarthCube might be--could exist without some notion of testing and integration; and how data and software in such a system could exist and be used without some notion of verification.
--Chaitan

CALL FOR FOCUS GROUP PARTICIPATION

Posted by Kim Patten on April 13, 2012 at 3:33pm in Governance
Reply by Kim Patten on April 16, 2012 at 12:07pm
All - my apologies, I shared the spreadsheet but did not make it editable. It is now open for editing for those with the link. Thanks to those who notified me on this issue.

Reply by Andrew Maffei on April 17, 2012 at 4:40pm
I just went to the Google spreadsheet and added another suggested focus group called "Framework for Vocabulary / Ontology (Semantic) Governance".
The idea is to consider whether a future EarthCube governance structure might (or might not!) include a governance model that could be made available to individual scientists and earth science projects to help them govern existing or proposed controlled vocabularies, ontology models, etc. used for support of cross-disciplinary science research. It's not clear to me if this "deserves" to be its own governance element of EarthCube or belongs elsewhere. I'd like to discuss this with others in the context of EarthCube governance. I'd be willing to chair this focus group if we can get a critical mass of participants for a session.
Reply by Geoffrey Fox on April 27, 2012 at 2:14pm
For Workforce (Education, Outreach, and Training; Underrepresented Populations) focus group, I note recent OCI report on this area http://www.nsf.gov/od/oci/taskforces/TaskForceReport_Learning.pdf

Reply by Bruce Caron on April 30, 2012 at 6:18pm
A really useful overview of IT management and how governance creates decision processes. Mapping these ideas onto the terrain of EarthCube data resources brings up some complications. I would think it might push EarthCube into a more federated model. Another issue that corporations tend not to have is how governance creates not just decision processes but also creates opportunities for emergent leadership and community ownership of these decisions and processes.

Reply by Genevieve Pearthree on April 30, 2012 at 7:10pm
Thanks for your comments. They are certainly insightful and bring up important points for consideration. Do you have any references/resources to recommend regarding governance of non-corporate entities, or that focus on the creation of emergent leadership and community ownership of decision-making, as you mentioned in your comment? I would like to go in that research direction but don't have a lot of resources to review as of yet. Any suggestions you might have would be very helpful.

Bruce Caron on May 2, 2012 at 4:12pm
Hi Genevieve,
Here are the references to a series of blogs I wrote on "double-loop" governance for virtual organizations:
http://cybersocialstructure.org/2012/04/23/double-loop-governance-references/
Jono Bacon's book on The Art of Community is particularly useful for the genre of community-led organization-of-organizations that EarthCube will likely be (or want to be).

Reply by Genevieve Pearthree on May 3, 2012 at 11:51am
Bruce,
Thanks for your help. I've added your references to my list of works to review. If other references come up, or if you have any other comments/ideas/suggestions, please feel free to get in touch.
Genevieve

REFERENCES

Abott, Mark. "A New Path for Science?" in *The Fourth Paradigm*, edited by Tony Hey, Stewart Tansley, and Kristin Tolle, 111-116. Redmond, Washington: Microsoft Research, 2009.

Advisory Committee for Cyberinfrastructure Task Force and Workforce Development, "Final Report." National Science Foundation, 2011.

Ahern, Tim. "A New Capability: Service Based Vertical and Horizontal Integration Platform." National Science Foundation EarthCube Expression of Interest 1041, 2012.

Ahern, Tim, and Keith Koper, "Flexibility in an EarthCube Design Approach: Ideas from IRIS and its Partners." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Ahern, Tim, David Simpson, and Brian Stump, "Guiding Principles and Best Practices from the IRIS Data Management System," National Science Foundation EarthCube White Paper: Designs Category, 2011.

Allison, M. Lee, "GeoNet: A Confederation of Scientific Data Network." National Science Foundation EarthCube Expression of Interest 1064, 2012.

Allison, M. Lee, Linda Gunderson, Kevin Gallagher, Stephen Richard, and Vivian Hutchison, "Geological Survey Contributions to Building Cyberinfrastructure in the Geosciences." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Allison, M. Lee, Gundersen, Linda C., Richard, Stephen M., and Dickinson, T.M. (2008). Implementation Plan for the Geosciences Information Network (expanded abstract), *In* Brady, S.R., Sinha, A.K., and Gundersen, L.C., eds. *Geoinformatics 2008—Data to Knowledge, Proceedings*: U.S. Geological Survey Scientific Investigations Report 2008-5172. pp. 9-11.

Arctur, David, "Expression of Interest in Development of an Organizational Framework for EarthCube." National Science Foundation EarthCube Expression of Interest 1053, 2012.

Arctur, David, Robert Downs, Robert Chen, Stefano Nativi, and David Schell, "Governance: An OGC White Paper for NSF EarthCube." National Science Foundation EarthCube White Paper: Governance Category, 2011.

Babaie, Hassan, "Semantic Web Technology-Driven Geoscience Network." National Science Foundation EarthCube White Paper: Designs Category, 2012.

Bedard, Dan, Ray Idaszak, Howard Lander, and Stan Ahalt, "What Open Source Software Development Teaches Us About Scientific Governance." National Science Foundation EarthCube White Paper: Governance Category, 2011.

Blanchard, Anita and Lynne Markus, "Sense of Virtual Community-Maintaining the Experience of Belonging," *Proceedings of the 35th Hawaii International Conference on System Sciences, Computer Society, Institute of Electrical and Electronics Engineers*, 2002.

Bowring, Jim. "Building a Software Engineering and Informatics Community of Practice and a Software Foundation for the Advancement of EarthCube." National Science Foundation EarthCube Expression of Interest 1022, 2012.

Bowring, Jim, Samuel Bowring, and Douglas Walker, "An Earth Cube Design Approach." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Braeckel, Aaron, Bob Barron, Arnaud Dumont, and Bruce Carmichael, "EarthCube Design Approach." National Science Foundation EarthCube White Paper: Designs Category, 2012.

Bristol, Sky, Roland Viger, Rich Signell, Nate Booth, Mike Frame, Linda Gundersen, and Cheryl Morris, " Role of Government Data Agencies, Virtual Observatories, and Communities of Practice in Sustaining EarthCube." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Broß, Justus, Harald Sack, and Christoph Meinel. "Encouraging Participation in Virtual Communities: The "IT-summit-blog" Case." *IADIS International Journal on WWWInternet*. 5 no. 2 (2007): 113-129.

Carlson, David, Shelley Olds, Steve Whitmeyer, Bruce Caron, Cathy Manduca, and John Tabor, "A Need for an Education and Outreach Component of EarthCube." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Caron, Bruce, "Earth Cube as a Double Loop Organization," *Virtual Democracy* (blog), April 4, 2012, http://cybersocialstructure.org/2012/04/04/earth-cube-as-a-double-loop-organization.

Caron, Bruce, "Realizing expectable returns on EarthCube investments in community building and democratic governance." National Science Foundation EarthCube White Paper: Governance Category, 2011.

Cosley, D., Frankowski, D., Kielser, S., Terveen, L., and Riedl, J., 2005, How Oversight Improves Member Maintained Communities, CommunityLab

Couch, Alva and Alex Bedig, "Building Plausible Digital Representations of Physical Entities." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Crane, Gary, "Designing a Community-Based Framework for Modeling the Consequences of Climate Change on Coastal Systems," National Science Foundation EarthCube Expression of Interest 1016, 2012.

Davenport, Thomas, Robert Eccles, and Laurence Prusak, "Information Politics." *Sloan Management Review* 34, (Fall 1992): 53-65.

Davis, Ethan, Douglas Dirks, Linda Miller, and Tom Yoksas, "An EarthCube Design Process: Unidata's Perspective." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Delaney, John and Roger Barga. "A 2020 Vision for Ocean Science?" edited by Tony Hey, Stewart Tansley, and Kristin Tolle, 27-38. Redmond, Washington: Microsoft Research, 2009.

DeLuca, Cecelia, "Commodity Governance for Communities of Communities." National Science Foundation EarthCube White Paper: Governance Category, 2011.

Downs, Robert, Robert Chen, and Mark Becker, "Data Management Training and Workforce Development for the Earth Science Community," National Science Foundation EarthCube Expression of Interest 1103, 2012.

Dyrud, Lars, Jonathan Fentzke, Robin Barnes, Brian Anderson, Larry Paxton, Elsayed Talaat, Daniel Morrison, Shin Oshanti, and Jesper Gjerloev, "A Future Vision In Action at APL for Cyberinfrastructure Enabled Geoscience." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Earth and Space Science Informatics Summit Report, March 13 – 14, 2008, Rome

Earth Science Information Partners Discovery Cluster (ESIP-DC), "A Lightweight Approach to Earth Science Data Discovery." National Science Foundation EarthCube White Paper: Designs Category, 2011.

Edwards, P.N., Jackson, S.J., Bowker, G.C., and Knobel, C.P. "Understanding Infrastructure: Dynamics, Tensions, and Design: Report of a Workshop on History & Theory of Infrastructure, Lessons for New Scientific Cyberinfrastructures." Arlington, Va.: 2007.

Fielding, Eric, Roweena Lohman, David Schmidt, Shimon Wdowinski, Falk Amelung, Rob Mellors, and Chuck Meertens, "Challenges Facing the WInSAR Consortium from the Coming Tsunami of SAR Observations." National Science Foundation EarthCube White Paper: Governance Category, 2011.

Fox, Geoffrey, "OpenQuake Social Networking for EarthCube." National Science Foundation EarthCube Expression of Interest 1055, 2012.

Fox, Geoffrey, Andrea Donnellan, Jay Parker, Marlon Pierce, Dennis McLeod, John Rundle, Lisa Grant Ludwig, and Chuck Meertens, " EarthCube and Earthquake Science." National Science Foundation EarthCube White Paper: Designs Category, 2011.
Fox, Peter, and Hendler, James. Semantic Escience: Encoding Meaning in Next-Generation Digitally Enhanced Science. *In* Hey, Tony, Tansley, Stewart, and Tolle, Kristin, editors. The Fourth Paradigm.147-152. Redmond, Washington: Microsoft Research, 2009.
Fulker, David, "Collaboration, Alignment and Leadership, invited piece in 'NDSl Reflections," an online series accessed at http://nsdlreflections.wordpress.com, 2008.
Fulker, David, 2011, "EarthCube Leadership." National Science Foundation EarthCube White Paper: Governance Category, 2011.
Gardler, Ross and Gabriel Hanganu, "Governance Models." Open Source Software Watch, last modified February 14, 2012. http://www.oss-watch.ac.uk/resources/governanceModels.xml.
Greenberg, DS, "Mohole: The Project That Went Awry (i)." *Science* 143, no. 3602 (1964): 115-119.
Greenberg, DS, "Mohole: The Project That Went Awry (ii)." *Science* 143, no. 3603 (1964): 223-227.
Greenberg, DS, "Mohole: The Project That Went Awry (iii)." *Science* 143, no. 3604 (1964): 334-337
Greenberg, DS, "NSF Appropriation: Mutiny on the Mohole." *Science* 152, no. 3724 (1966): 895-896
Gundersen, Linda, C. (2008). Answers to Earth Systems Science Questions: The evolution of Geoinformatics at the U.S. Geological Survey. *In* Brady, S.R., Sinha, A.K., and Gundersen, L.C., eds. *Geoinformatics 2008—Data to Knowledge, Proceedings*: U.S. Geological Survey Scientific Investigations Report 2008-5172. pp. 2-3.
Hanganu, Gabriel and Ross Gardler, "Planning for Sustainability," Open Source Software Advisory Service, last modified December 12, 2011, http://www.oss-watch.ac.uk/resources/planningsustainability.xml
Hoffman, Michael, Judith Curry, and Calton Pu, "EarthAgora: A Collaborative Knowledge Management Tool." National Science Foundation EarthCube White Paper: Designs Category, 2011.
Holt, William, and Mary Miller, "Community Governance: UNAVCO Facility Models for EarthCube Implementation." National Science Foundation EarthCube White Paper: Governance Category, 2011.
Idaszak, Ray (ed.), Stan Ahalt, Barbara Minsker, Dave Tarboton, Michael Tiemann, James Brunt, Robert Tawa, Peter Backlund, Mark Williams, and Mark Parsons. "Insights and Precepts from a Workshop on Creating a Software Institute for Environmental Observatories." National Science Foundation EarthCube White Paper: Governance Category, 2011.
Kelbert, Anna, and Gary Pavlis, "Earth Sciences: The Tower of Babel." National Science Foundation EarthCube White Paper: Designs Category, 2011.
Kumar, Shailendra, Stefan Falke, Martin Frederick, and Rick Ohlemacher, "A Public-Private Partnership Opportunity for NSF Earth Cube to Serve Stakeholders and Decision Makers National Science Foundation EarthCube White Paper: Governance Category, 2011.
Leetaru, Hannes, "EarthCube – Strategic Organizational Framework (Centralized Model)." National Science Foundation EarthCube Expression of Interest 1015, 2012.
Leetaru, Hannes, Michael Welge, Bernie Ács, Shaowen Wang, Yu-Feng Forrest Lin, and Kenton McHenry, "EarthCube Design Approach." National Science Foundation EarthCube White Paper: Designs Category, 2011.
Leetaru, Kalev, "A Decade and a Half of Archiving the Web for Data Mining: Lessons Learned and How Users Use Web Archives," International Internet Preservation Consortium 2012 General Assembly. Library of Congress, Washington DC, April 30-May 4, 2012.

Lenhardt, Christopher, “Governance or Process: Building Community Through Collaboration for Shared Solutions.” National Science Foundation EarthCube White Paper: Governance Category, 2011.
Long, James, and Karen Remick, “A Data-Based Professional Networking System.” National Science Foundation EarthCube White Paper: Designs Category, 2011.
Mader, Christopher, Benjamin Kirtman, Felimon Gayanilo, Joel Zysman, and Nicholas Tsinoremas, “An Architectural Approach for Sharing, Discovery and Knowledge Dissemination on EarthCube, National Science Foundation EarthCube White Paper: Designs Category, 2011.
Maechling, Philip, "Focus on Forecasts.” National Science Foundation EarthCube White Paper: Designs Category, 2011.
Maidment, David, “EarthCube – An Open Geoscience Community.” National Science Foundation EarthCube Expression of Interest 1019, 2012.
Malone, Thomas, Robert Laubacher and Chrysanthos Dellarocas, “Harnessing Crowds: Mapping the Genome of Collective Intelligence,” MIT Center for Collective Intelligence, 2009.
Michener, William, Rebecca Koskela, Matthew Jones, Robert Cook, Mike Frame, Bruce Wilson, and David Vieglais, “DataONE-Enabling Cyberinfrastructure for the Biological, Environmental and Earth Sciences.” National Science Foundation EarthCube White Paper: Designs Category, 2011.
National Science Foundation, Advisory Committee for Geosciences. “GEO Vision Report.” October 2009.
Nestor, Stilpon, “International Efforts to Improve Corporate Governance: Why and How,” Organization for Economic Cooperation and Development, 2001.
Orcutt, John, Kate Keahey, Matthew Arrott, Bill Howe and Oscar Schofield, “Community-Based Earth Observatories: Implementation and Integration.” National Science Foundation EarthCube White Paper: Designs Category, 2011.
Percivall, George (ed.), Amy Apon, Lorenzo Bigagli, Jack Cothren, Philip Dibner, Gerd Heber, Craig Lee, Fred Limp, Yon Liu, Jay Pearlman, Sven Schade, Ingo Simonis, Don Sullivan, Dawn Wright, and Yan, “Cyberarchitecture for the Geosciences.” National Science Foundation EarthCube White Paper: Designs Category, 2011.
Peters, Shanan, Miron Livny, Kurt Squire, Jack Williams, Bill Hibbard, Puneet Kishor, Sam Batzli, Patrick McLaughlin Susan Millar, Tom Whittaker, Nancy Wiegand, and Terry Millar, “Towards a New Distributed Platform for Integrative Geoscience: An EarthCube Design Approach and Prototype Plan.” National Science Foundation EarthCube White Paper: Designs Category, 2011.
Ramamurthy, Mohan. “Unidata Governance: A Quarter Century of Experience.” National Science Foundation EarthCube White Paper: Governance Category, 2011.
Ramamurthy, Mohan, “Governance Examples” (presentation, EarthCube Governance Virtual Plenary Workshop), April 11 and 17, 2012.
Richard, Stephen, and Vivian Hutchison. “Recommendations for the Future of the U.S. Geoscience Information Network.” Arizona Geological Survey, 2011.
Stump, Brian, and David Simpson, “The IRIS Consortium Structure and Governance,” National Science Foundation EarthCube White Paper: Governance Category, 2011.
Tarboton, David, David Maidment, Ilya Zaslavsky, Jon Goodall, Dan Ames, Richard Hooper, Jeffrey Horsburgh, Kim Schreuders, Ray Idaszak, and Alva Couch, “Advancing Solutions for an EarthCube Design. What can be learned from the CUAHSI HIS experience? National Science Foundation EarthCube White Paper: Designs Category, 2011.

Tatham, Elizabeth. “Roles in Open Source Projects,” Open Source Software Advisory Service, last modified May 15, 2012, http://www.oss-watch.ac.uk/resources/rolesinopensource.xml

Walker, J. Douglas, Allison, M. Lee, and Gundersen, Linda, convenors. Workshop on Working towards a National Geoinformatics Community (NGC), USGS Denver Federal Center, Denver, Colorado, September 23-24, 2010.

Weill, Peter, and Jeanne Ross. *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*. Boston: Harvard Business School Press, 2004.

Wenger, Etienne, Richard McDermott, and William Snyder. *Cultivating Communities of Practice.* Boston: Harvard Business School Press, 2002.

Wilbanks, John. “I Have Seen the Paradigm Shift, and It Is Us.” In *The Fourth Paradigm*, edited by Tony Hey, Stewart Tansley, and Kristin Tolle, 209-214. Redmond, Washington: Microsoft Research, 2009.

Zaslavksy, Ilya, “Readiness of Disciplinary Data Systems for Cross-Domain Interoperability within a Standards-Based EarthCube Reference Framework,” National Science Foundation EarthCube Expression of Interest 1051, 2012.

Zaslavksy, Ilya, Mark Williams, Anthony Aufdenkampe, Kerstin Lehnert, Emilio Mayorga, and Jeff Horsburgh, “Data Infrastructure for the Critical Zone Observatories (CZOData): an EarthCube Design Prototype.” National Science Foundation EarthCube White Paper: Designs Category, 2011.