EarthCube: A transformational concept for the geosciences

CLIFFORD A JACOBS
SENIOR ADVISOR FOR THE GEO SCIENCES
NATIONAL SCIENCE FOUNDATION
USA
# Acknowledgements

## Leadership

- Tim Killeen, Assistant Director for Geosciences
- Alan Blatecky, Head of Office of Cyberinfrastructure

## EarthCube Team

<table>
<thead>
<tr>
<th>Eva Zanzerkia (GEO)</th>
<th>Mark Suskin (OCI)</th>
<th>Dane Skow (OCI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbara Ransom (GEO)</td>
<td>Irene Lombardo (OCI)</td>
<td>Robert Chadduck (OCI)</td>
</tr>
<tr>
<td>Lisa Boush (GEO)</td>
<td>Jennifer Schopf (GEO)</td>
<td>Almadena Chtchelkanova (CISE)</td>
</tr>
<tr>
<td>Rosalind Douglas (GEO)</td>
<td>Melissa Lane (GEO)</td>
<td></td>
</tr>
</tbody>
</table>
## Context
- Major Science and Cyber-Drivers

## The Vision
- Transformational

## Activities to date
- Discovery about the process, appreciation for the social aspects, and exceeded expectations

## Balance & Challenges
- Broadening participation
- Understanding the social context

## Reflections
- Rapid Prototyping of the process

## Next Steps
- Facilitate a synergic set of activities

## Questions/comments
CONTEXT -------- SCIENCE AND CYBERINFRASTRUCTURE

EarthCube
Science Context

“FOSTERING A SUSTAINABLE FUTURE THROUGH A BETTER UNDERSTANDING OF OUR COMPLEX AND CHANGING PLANET.”

NSF’S GEO VISION REPORT, 2009
PURPOSE: “To understand more deeply the planet and its interactions will require the geosciences to take an increasingly holistic approach, exploring knowledge coming from all scientific and engineering disciplines.”

CALL TO ACTION: “Over the next decade, the geosciences community commits to developing a framework to understand and predict responses of the Earth as a system—from the space-atmosphere boundary to the core, including the influences of humans and ecosystems.”
Era of Observation and Simulation

Arctic Sea Ice
Research Vessel Sikuliaq

Oceans

EarthScope Observatory Network

Water

Satellites

Earth System Modeling
The Data Pyramid
(geosciences research data exists in every part of the pyramid)

The data pyramid – a hierarchy of rising value and permanence
Adapted from Francine Berman article in Communications of the ACM
Cyberinfrastructure Context

“…..TO BUILD NEW TYPES OF SCIENTIFIC AND ENGINEERING KNOWLEDGE ENVIRONMENTS AND ORGANIZATIONS AND TO PURSUE RESEARCH IN NEW WAYS AND WITH INCREASED EFFICACY.”

REVOLUTIONIZING SCIENCE AND ENGINEERING THROUGH CYBERINFRASTRUCTURE: REPORT OF THE NATIONAL SCIENCE FOUNDATION BLUE RIBBON ADVISORY PANEL ON CYBERINFRASTRUCTURE, 2003
Science and Society Transformed by Data

- **Modern geoscience**
  - Data- and compute-intensive
  - Integrative, multi-scale

- **Multi-disciplinary collaborations to address complexity**
  - Individuals, groups, teams, communities

- **Sea of Data**
  - Age of Observation
  - Distributed, central repositories, sensor-driven, diverse, etc
Cyberinfrastructure Ecosystem
(Cyberinfrastructure for 21st Century)

- Maintainability
- Sustainability
- Extensibility
The EarthCube Vision

TRANSFORM THE CONDUCT OF RESEARCH BY FOCUSING ON OUTCOMES AND PROCESS
EarthCube Vision
focused on Outcomes & the process to get there

Goal

To transform the conduct of research in geosciences by supporting the development of community-guided cyberinfrastructure to integrate data and information across the Geosciences.

Outcomes

1. Transform practices within the geosciences community over the next decade
2. Provide unprecedented new capabilities to researchers and educators
3. Vastly improve the productivity of community
4. Accelerate research on the Earth system
5. Provide a knowledge environment framework for the geosciences
### Elements of a knowledge environment

#### Making Knowledge Visible
**Easy Usability**
- Who knows what
- Taxonomy of expertise
- Yellow pages
- Competence

#### Building Knowledge Intensity
**Creation (local)**
- Training, face-to-face contacts
- Competence centers
- Community of practices
- Management of knowledge processes
- Networking

#### Building Knowledge Infrastructure
**Global Access**
- Common communication infrastructure
- Access to external/internal information/knowledge/sources
- Use of modern methods and tools

#### Developing a Knowledge Culture
**Motivation Enablers**
- Value and culture
- Rewarding
- Sharing/exchange of knowledge
- Shared mindsets and visions
- Trust of each other

---

*From A. Boynton, “Exploring Opportunities in Knowledge Management: How to Get Started” Knowledge Management Symposium: Leveraging Knowledge for Business Impact, IBM Consulting, Sydney, 1996*
An alternative approach to respond to daunting science and CI challenges

EarthCube is an outcome and a process

EarthCube will require broad community participation

The unlabeled dots represent the “long-tail” of science which is graphically under represented in this diagram. Also, the large dots under represent all the community-guided activities in support of geosciences.
Given: almost all the technologies used today to provide cyberinfrastructure to the geosciences will be refreshed in the next decade.
Some Challenges to Overcome

- Collection
- Trust
- Usability
- Interoperability
- Diversity
- Security
- Education and training
- Data publication and access
- Commercial exploitation
- New social paradigms
- Preservation and Sustainability

Riding the wave
How Europe can gain from the rising tide of scientific data

Final report of the High Level Expert Group on Scientific Data A submission to the European Commission October 2010
EarthCube Activities to date

• EARLY ENGAGEMENT OF SOCIAL SCIENCE
• EARLY ENGAGEMENT OF CI SPECIALISTS NECESSARY, BUT NOT SUFFICIENT
Institutional and systems requirements

• Creating Value . . . expanding the “pie” and enabling systems transformation

• Mitigating Harm . . . anticipating and mitigating externalities and catastrophic systems failures

• Stability & Agility . . . Fair, effective, robust “rules of the game”
EarthCube Stakeholder Alignment

- Early EarthCube award
- Principal Investigator
  - Joel Cutcher-Gershenfeld, University of Illinois, Urbana-Champaign
  - Experience with understanding the dynamics of stakeholder alignment
- Survey
  - Earlier EarthCube participants
  - Broad geoscience researchers not initially involved with EarthCube
- Research in visualizing stakeholder alignment
  - Developed the Z-Flowers representation
Using “Z-Flowers™” to view the stakeholder survey results

When seeing every stakeholder’s perspective matters, the Z-Flower™ are a new, powerful way of visually examining the distribution of views. Each hexagon represents a different stakeholder and is color coded based on his or her response to a 17 point “visual analog scale” such as the one on the right. You “read” a Z-Flower™ by viewing the color pattern – to see how positive or negative it is overall, if there are “pockets” that are at the extremes (positive or negative), or other patterns. The key is that you can see each stakeholder’s perspective on a given issue.

1. The center hexagon signals the mean.

2. The hexagons (one for each stakeholder) are tiled in a spiral out from the middle beginning with the responses closest to the mean – so the outliers are on the outer edges. Positive or negative views may alternate based on what is closest to the mean.

3. Missing or NA responses.

4. Very negative or positive outliers.

Three illustrative Z-Flowers™
Importance and ease of access within field/discipline

Comment: Vast majority selecting 16 out of 16 on importance; dozens reporting barriers on ease of access, but some bright spots.

<table>
<thead>
<tr>
<th>How IMPORTANT is it for you to find, access, and/or integrate multiple datasets, observations, visualization tools, and/or models in your field or discipline?</th>
<th>High $\mu$ (α)</th>
<th>Low $\mu$ (α)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.0 (2.0)</td>
<td>13.8 (3.3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How EASY is it for you to find, access, and/or integrate multiple datasets, observations, visualization tools, and/or models in your field or discipline?</th>
<th>High $\mu$ (α)</th>
<th>Low $\mu$ (α)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4 (4.0)</td>
<td>6.8 (3.9)</td>
<td></td>
</tr>
</tbody>
</table>
Social Network Established Early

http://earthcube.ning.com/

• ~1200 members to the EarthCube website
• 113 white paper submission; 185 respondents to user survey
• ~70 expression of interest emails
• 27 Groups
• Unknown number of hours of pro bono contributions by the community
• Unprecedented view of the pulse of the geosciences community
2011
- Accelerating the Community Dialog
- Defining the initial scope of EarthCube
- New starting point for collaboration

2012
- Developing convergence/consensus
- Forming recommendations for development
- Examining governance

First DCL
- 6/11
- 7/11
- 11/11
- 2/12
- 6/12

EAGER Phase start
- 6/12
- 7/12
- 8/12
- 9/12
- 10/12

Road maps

Road Map updates

Webex Community Outreach

First Charrette

Second Charrette

First PI Workshop

Second PI Workshop

Europe Workshop
Collaboratively produced framework to form an integrated & synergistic path forward

Community Event
Charrette 2

- Data Discovery/Mining/Access
- Semantics and Ontologies
- Workflow
- Governance

- Hydrosheric Model (OHMF)
- Data GeoData
- Data Brokering
- X-Domain Interop.
- Service Based Integration
- Layered Architecture
- Earth System Modeling
Roadmaps

- Community Groups
  - Data Discovery, Mining, and Access
  - Governance
  - Semantics and Ontologies
  - Workflow
- Concept Projects
  - Brokering
  - Cross-Domain Interoperability
  - Earth System Modeling
  - Layered Architecture
  - Web Services

- Purpose
- Communication
- Challenges
- Requirements
- Status
- Solutions
- Process
- Timeline
- Management
- Risk

http://earthcube.ning.com/page/draft-roadmap
Maintaining Balance & Appreciating the Challenges

BROADENING THE DIALOG AND FOCUSING THE EFFORT
Balance needs to be adjusted to reflect user needs and aspirations.

**Cyberinfrastructure Elements “on the table:”**
- Data discovery, mining and integration
- Semantics and ontologies
- Workflow
- Brokering
- Cross-domain interoperability
- Dark geodata
- Earth system model
- Layered architecture
- Modeling framework
- Web services
- Physical samples
- Software engineering
- Website and collaboration environment

**Additional Social/Institutional Elements:**
- Norms for sharing data/models
- Established communications patterns
- Relations between Cyber and Geo communities
- Professional associations
- Promotion and tenure systems
- Roles of government agencies and private industry
- International partnerships

**Social Elements & End-User Needs Not addressed**
- Governance
- Discipline-specific needs and drivers
- Education and workforce development
- Collaboration environment

Adapted from presentation by Joel Cutcher-Gershenfeld, University of Illinois at Urbana-Champaign.
NSF engages researchers, educators, and social scientists

- Early engagement of social scientist to assess stakeholder alignment
  - Valuable insight
  - A basis for assessing EarthCube’s effectiveness
- Encourage many discipline-specific workshops to engage leading researchers and educators
  - Special emphasis of early career scientists
  - Address issues within a consistent framework
- Must consider motivations (technical and social) to get the most out of community engagement
Motivations: technical

- Getting Science done now and in the future
  - Science drivers and aspirations
- Identifying and addressing barriers and challenges
  - Increase productivity and capability of researchers and educators
- Assessing the distribute of resources (data and CI) and access to them
  - EarthCube must address the needs of big, medium, and small science activities.
- Understanding the multifaceted and multivariate conduct of research within the geosciences
  - Current use cases and hoped for future capabilities
Motivations: social

Motivating questions:

• Will geoscientists share data, models, tools, and visualizations – if you build it, will they come?
• Will geoscientists and cyber/computer scientists collaborate to advance EarthCube – what are the incentives to cooperate?
• Will interdisciplinary work occur on a sufficient scale to tackle pressing earth systems research challenges – is there a clear, shared success vision?

Top Six Barriers to Sharing Data (survey):

• No time/Not enough time for QA/QC
• No repository or known repository
• Inadequate standards, standardized formats, etc.
• Want to publish first/not be scooped
• File size too large/server size too small
• No credit/incentive for sharing
Additional Consideration

- Variation by fields and disciplines
- Variation by age
- Variation by home institution
- Variation in social systems associated with data sources
- Desirability vs. feasibility
- Accelerating technological change
Reflections

WHAT WE LEARNED SO FAR AND WHAT WE WORRY ABOUT ➔ RAPID PROTOTYPING THE PROCESS
NSF’s Seven Modes of Success

We are proactive

We began with the end goal in mind

We prioritized EarthCube tasks

We emphasized a non-competitive & broadly inclusive process

We listen to the community and shared our vision for EarthCube

We facilitated synergy within and across communities

We engaged and energized NSF colleagues in the process
Seven Modes of Failure

1. Unrealistic or misaligned expectations among people presently involved in EarthCube
2. “Build it and they will come” mindset
3. Not valuing what presently exists
4. Not advancing the frontier in transformative ways
5. Not engaging the 14,000+ geoscience and cyber stakeholders
6. Not anticipating the needs of the next generation of geoscience & cyber stakeholders
7. “Unk Unk” – additional unknown unknowns
Questions

What is the best way to take advantage of the current state of EarthCube?

What portfolio of activities are needed to engage the scientific community in the development/use of EarthCube?

What are effectively mechanisms to engage our sister agencies and international partners in the EarthCube dialog?

What are effective ways NSF can facilitate collaborative dialogs among the geosciences?

What does it mean to change the community culture?
Next Steps

Engage broadly in creating the building blocks of EarthCube
More Engagement and Diversity

- EarthCube solicitation – soon
  - One solicitation with amendments
- Establishing Governance
- Community Engaging Workshops
- Developing an architecture
- Leveraging existing efforts
- Further engagement of Federal agencies, International partner, and the private sector
Questions and/or Comments

EARTHCUBE@NSF.GOV  CJACOBS@NSF.GOV