

## EarthCube Working Group Proposal

1. **Working Group name:** Technical Requirements and Synthesis from Science Use Cases
2. **Submission date:** Initial submission 20 November, 2014. Revised submission December 5, 2014
3. **Short name (no more than 3 words):** TAC Use Cases
4. **Statement of need and work (500 words max):**

### Statement of Need

Between 2012 and 2014, EarthCube conducted a series of End-User Workshops designed to gather scientific challenges and cyberinfrastructure needs from NSF Geoscience domains and related stakeholders. Output from these workshops contains a collection of science use cases, illustrating domain-specific science challenges, and the technical capabilities needed to address those challenges. The workshops are part of a continuing emphasis on a user-driven development of EarthCube, with additional use cases produced by other EarthCube activities, such as the building block projects.

These use cases provide detailed information about the types of research that may be conducted in the near future and the current cyberinfrastructure barriers to achieving research goals. These needs and barriers represent gaps in the current cyberinfrastructure landscape in support of discipline-specific or interdisciplinary research and provide a valuable foundation for planning the EarthCube cyberinfrastructure system needed by the Geoscience community. Therefore, there is a critical need for this information to be compiled and synthesized from their various sources into a coherent document that can inform EarthCube architecture planning of the Geoscience community's technical requirements. In doing this, the working group will reach out to other EarthCube organizations and activities for both inputs and coordination.

The working group deliverables are designed to be a persistent resource available to the TAC and TAC working groups as they strategize and prioritize next steps. However, we expect the outcomes to have broader value as public EarthCube resources. CI developers will find motivating scenarios and priorities, scientists will find a way to understand how the technology will serve their needs, and EarthCube leadership and outreach will see explicit connections between CI efforts and science requirements.

### Statement of Work

This work will be completed in three phases.

- 1) **Identification of Sources:** First, the existing sources of use cases within EarthCube will be identified.
  - a) Workshop reports are already written and available from the EarthCube workspace.
  - b) Funded projects will be contacted to request use cases they have developed (this will be coordinated TAC's planned funded project survey).

- c) We will also request assistance from the TAC, SSC and Office in identifying any other critical sources of use cases in EarthCube (with the caveat that a thorough review of the entire set of EarthCube artifacts, including white papers and roadmaps, is likely not tractable.) This could include recommendations of the EarthCube roadmaps developed as part of the EAGER awards two years ago.

- 2) **Summarizing Use Cases:** The second phase will be reviewing the documents to extract and summarize the use cases in a structured form. Part of this work will be defining an appropriate structure and format for the summaries. They should be brief while still capturing essential characteristics, allow categorization based on several technical and domain facets, and be easily searched and browsed. While it is beyond the scope of this project to develop a tailored tool, the summaries may require a simple database. A decision about the appropriate format of the use cases will be made in consultation with other EarthCube committees and working groups and the Office. We will request liaisons from the TAC working groups on gap assessment and test bed formulation to help ensure that the use case are structured to contribute to their efforts as fully as possible, and one or more liaisons from the SSC to provide coordination with their efforts and scientific perspective (several members of the proposed working group already participate in the SSC).

We will reach out to the organizers of each end-user workshop, and the leads of each funded project, to request a volunteer to lead or assist with the summarization of use cases from their workshop or project. This will ensure that the use cases summaries remain scientifically accurate, as well as bring additional people to the task. We will also ask the Office to circulate a request for participation from the broader EarthCube community (e.g. through a newsletter item or a website notice.)

- 3) **Synthesizing Use Cases:** The third and final stage will be a synthesis of the summaries. This synthesis will include:
  - a compilation of technical requirements within and across end-user communities represented in the EarthCube workshops and funded projects, highlighting the breadth of needs, identification of substantial barriers, and relative priorities;
  - the identification of gaps in the coverage of existing use cases (coverage will be assessed from the perspective of cyberinfrastructure requirements, and not domain science coverage, which is the purview of the Science Committee);
  - the identification of commonalities across use cases. When feasible, meta use cases will be developed to express requirements that cut across multiple domain-specific use cases. For example, scientists studying the behavior of earthquakes, volcanoes, and hurricanes may have similar workflow needs even though the fundamental data and calculations are different.
  - An assessment of the future value of the use case summaries. Based on feedback from the TAC and other EarthCube governance bodies, is this an effort that has one-time value during the current planning stage, or a

resource that should be maintained and expanded through time as EarthCube activities continue to articulate use cases?

We will liaise with the TAC Gap Analysis Working Group, in addition to reporting to the TAC as a whole, to ensure that our use case gap analysis is useful to their gap analysis effort. We will also work with the proposed Testbed working group to identify representative use cases that can help in the formulation and assessment of an EarthCube testbed.

## **5. End goal and deliverables:**

The overall end goal is to review the substantial body of use case information that already exists in EarthCube artifacts, and to extract and summarize it in a format that is useful for future EarthCube architecture and technology planning. The two primary deliverables are:

1. structured summaries of the key use cases that have been collected to date through the EarthCube process; and
2. a synthesis and analysis of the use case summaries (described above.)

## **6. Committed participants (at least 3 and at least from 3 different organizations):**

- Karen Stocks, Scripps Institution of Oceanography
- Danie Kinkade, Woods Hole Oceanographic Institution
- Lisa Kempler, MathWorks
- Yolanda Gil, Information Sciences Institute, University of Southern California
- Corinna Gries, Center for Limnology, University of Wisconsin, Madison

## **7. Chairs (specify their time commitment to the working group):**

Karen Stocks, committed to at least 60 hrs of work total  
Danie Kinkade, committed to at least 40 hrs of work total

## **8. Final deliverables (with dates for draft releases and public requests for comments):**

January

- Draft inventory of documents to be assessed for use cases. Input will be requested from the TAC, the SSC, and the Office (initiated at an in-person meeting at AGU in Dec.)

February

- Inventory of documents to be assessed for use cases revised and prioritized
- First draft of a template (format and structure) for use cases developed, circulated for comment to TAC and TAC working groups, and SSC liaison.

March

- 10 use cases summarized using template; template revised based on experience.

April

- Use case summaries complete

May

- Initial draft of synthesis of a set of use cases for one domain, e.g. oceanography, atmospheric sciences (acting as a model for the larger synthesis)
- Any edits to use cases prompted by synthesis effort

June

- Draft synthesis technical requirements within and across use cases, with public input requested at the 2015 All-Hands meeting (and through emails)

July

- comment period open

September

- delivery of the final summaries and synthesis to TAC.

**9. Timeline (no more than one year, with concrete goals for the first 3 months and planned goals for the rest of the quarters):**

9 months: January 2014 to September 2015

**10. Alignment with EarthCube goals and/or Standing Committee priorities:**

This working group proposal supports major functions of the Technical and Architecture Committee (as outlined in <http://workspace.earthcube.org/technology-committee>), specifically:

- Ensure the explicit connection between the scientific process and technical function, including coordinating testbeds and other mechanisms for development of cyberinfrastructure components, for use in science use cases
- Seek alignment of EarthCube funded projects to foster integrated technology to meet end user requirements and create new capabilities to enable transformative geoscience research
- Identify gaps in coverage of needed cyberinfrastructure capabilities, and determine recommendations on how to fill them.

**11. Risk assessment (what might lead to failure, how to mitigate those risks):**

Crucial to the task of consolidation and synthesis of domain-specific use case information is the need to coordinate with the Science Standing Committee (SSC) and its planned working group on use cases. Such coordination is necessary to ensure proper interpretation of the scientific needs for cyberinfrastructure technologies and services. Close communication or collaboration with SSC members can ensure this takes place throughout this project. This working group is committed to communicating with the SSC and providing collaborative opportunities for SSC members to engage in this effort.

Ensuring timely deliverables from a volunteer effort is always a risk. These risks will be mitigated by 1) committed time from the chairs; 2) outreach for additional participation; and 3) prioritization of the key sources of use cases. EarthCube has produced many documents that may contain use cases. The workshop reports and funded projects will be the initial targets for extracting use cases, with white papers, roadmaps, charette reports, etc. addressed only if time allows.

Part of the efforts include contacts with User workshop leaders and Building Block PIs. Experience has shown that simple emails do not elicit strong response. For those areas where such contacts are important, the WG will coordinate with the other TAC groups in formulating a questionnaire for funded projects and make individual contacts with key individuals as necessary.

### **12. Resources available to the working group (eg, datasets, people, IT, etc):**

The members of the use case working group have committed time to the project. The EarthCube End User Workshop reports and other documents are available through the EarthCube document repository on workspace.

### **13. Requested total estimated budget and brief budget justification:**

We request travel support for 1 member of the working group to attend the 2015 All-Hands meeting, to present results and request feedback on the draft use case synthesis. \$2000.

Travel funds for one face-to-face meeting of the working group to launch the synthesis effort. Estimated as 2 meeting days plus travel time for 5 people. The date and location is TBD, but the RDA meeting in San Diego in March is well-timed. \$7500

TOTAL REQUEST: \$9500

We also request the following non-financial support from the EarthCube Office:

- webex support or similar for hosting teleconferences
- administrative support for keeping email lists, scheduling meetings, sending meeting announcements, taking and posting meeting minutes.
- workspace or similar site for online collaboration: e.g., hosting documents, having discussion threads, keeping a group calendar.
- depending on the outcome of the use case summary design effort, we may request the Office to support an open-source database for group use.

### **14. Initially appointed host (the host will be committed to respond to inquiries about the group and will put on the group's web site a clear description of status):**

Karen Stocks and Danie Kinkade (rotating)