

# Response to the EarthCube Advisory Committee Report

*Delivered by the EarthCube Leadership Council*

*November 23, 2016*

## Introduction

EarthCube is a forward-looking program of the US National Science Foundation that was established to meet the aspirational research needs of the broad geoscience community through advanced cyberinfrastructure. EarthCube intends to develop facile solutions to far-reaching scientific objectives, exploiting rapidly growing data resources and emerging computational capabilities to transform the conduct of geoscience. EarthCube's vision is to foster modern modes of scientific inquiry that infuse technological innovations into the emerging practices of trans-disciplinary knowledge discovery and utilization, digital scholarship, data science and analytics, data and software stewardship, and open science, and basic research practice in geosciences.

EarthCube's progress over the first 4 years was reviewed in early 2016 by a Reverse Site Visit (RSV) Advisory Committee (AC), which summarized its findings and provided excellent recommendations for improvement and the future evolution of the program. [\[Link to RSV Report\]](#). The RSV committee's report focused on four broad areas: reviewing the Strategic Vision Document, the Roadmap Document, community engagement, and suggestions for next steps towards program sustainability. The report identified fundamental concerns with the trajectory of the program and gave detailed recommendations for future activities.

This official report describes the actions that have been or currently being implemented by the EarthCube Leadership Council (LC) and the EarthCube community in response to the RSV report, and aims to convey to the National Science Foundation (NSF) and the newly formed EarthCube Advisory Group (AG) the progress made over the nine months following the RSV and the trajectory for the future. In addressing the RSV report, the LC planned and then initiated actions to effect change across the EarthCube program, as described herein. The LC identified several issues as high priority and addressed these immediately, while targeting others for action in the subsequent months.

The work of EarthCube has indeed proven to be challenging so far, but the goals are worthwhile and likely unattainable by other means. Through this document the LC wishes to convey that these goals are achievable by the rapidly developing participant community with adaptive responses to stakeholder needs and advisory recommendations.

## Acknowledgement

The LC thanks the RSV AC for the in-depth review of the EarthCube program, and appreciates the insightful findings and recommendations laid out in their RSV report.

---

## LC Response to AC Findings and Recommendations

This response document, prepared by the LC, includes discussion of the current state of affairs and proposed programmatic modifications moving forward. After thorough consideration and exploration of paths forward, the LC resolved to take necessary steps as far as its mandate and decision-making authority to address the concerns expressed in the RSV report. These include targeted community actions as well as recommendations to the NSF for actions that require NSF approval and implementation.

**Response Structure:** Snippets of comments and recommendations from the AC report are included in blue text below for context. Refer to the [RSV Report](#) for the full text from the AC. The EarthCube LC responses follow each Recommendation.

### 1. GENERAL CONCERNS

*The AC notes that the LC has demonstrated an exceptional level of volunteer commitment to the success of NSF's EarthCube program. We recognize the cultural, social, and scientific challenges faced by the LC in attempting to promote a grassroots movement enabling the geoscience community to move towards a new data-enabled cross-disciplinary research paradigm. Over the past several years, steps have been taken to establish a governance model, populate the EC committees, and begin the work for EarthCube. The large number of community workshops and efforts to build consensus on community goals has clearly strained available intellectual and time resources. In weighing these struggles against accomplishments from the first 5 years, **it is evident that significant structural changes are needed to create a thriving sustainable program within the next 5 years.***

*The AC identified three interrelated problems impeding progress in the EarthCube initiative. The first is the persistent lack of clear definition of "EarthCube." Given the time and financial resources invested, there should be no ambiguity about program definition or objectives at this time. The second impediment is the reliance on a largely voluntary LC and a fully voluntary governance structure. Given the lack of clear program definition, self-selected stakeholders are free to tailor their understanding of EarthCube as they see fit, precluding the development of a common vision and common milestones. The third impediment is the lack of interconnection between LC activities, proposal solicitations, and assessment of outcomes from funded projects. In essence, the LC and the EarthCube PIs appear to be operating "open loop," except where interactions are mediated by the NSF. This, again, is a severe impediment to progress toward a common vision.*

*Based on these concerns, the AC recommends that a full EarthCube program review be conducted, which may result in a revamping of program objectives, governance structures, and implementation plan. More information on these concerns can be found in the detailed findings and recommendations presented in Sections 2 through 6.*

## **LC Response**

The LC agrees with the three overall concerns regarding Governance structure and reliance on volunteers, incomplete community alignment with a single working definition of EarthCube, and the functional disconnects between community goals, funded activities, solicitations, and ongoing assessment. Each of these concerns is discussed in the sections below, numbered and ordered the same way as the RSV report to make the responses and actions easy to follow by readers of this report.

## **2. STRATEGIC VISION**

*Charge: How well does the strategic vision represent the scientific and technological goals of the geosciences community? Does it present an appropriate framework to guide future EarthCube community activities?*

### **2.1. Finding: The definition of EarthCube remains unclear to the GEO community**

#### **Recommendation 1**

*Earthcube solicitations should be written to attract a broad base of innovative thinkers that cross the disciplines of data science and geoscience. EarthCube is a program within CIF21, and yet the data science component of the program is underemphasized. Solicitations should also be written to target the challenge of interoperability across traditionally disparate data collection methods and repositories. The first solicitation category should invite “blue sky” thinking, the second requires decisions that may exclude some players.*

## **LC Response and Actions Taken**

- 1. Emphasize EarthCube’s role in Data Science:** As a leader in Data Science thinking, EarthCube has built and continues to grow a broad community of data science experts that advances Data Science as a multi-disciplinary approach, involving individuals, projects, and facilities that bring innovation to the use of informatics in the Earth Sciences. These Data Science activities are not emphasized or highlighted sufficiently in current EarthCube communications. To address this issue, the LC recommended that the EarthCube Science Support Office (ESSO) add a ‘Data Science’ section to the EC website and consider data science highlights as part of communications activities.
- 2. Encourage Data Science in future solicitations:** The LC agreed that future recommendations to NSF solicitations will include language that address not only building blocks but also innovative approaches to data science.

3. **Interoperability as a Funded Project Requirement:** Interoperability is addressed within the [Funded Project Requirements document](#) that was generated by a Tiger Team established by the LC in response to the AC report. The report was provided to the NSF in June of 2016 in time for the new 2016 awards announcement and the requirements were provided to the Principal Investigators of the 2016 awards. This document addressed Interoperability through a set of Development Requirements, which included the use of EC-recommended standards, and standard reporting, development of modular APIs, and open data (and software) release with appropriate metadata. While these recommendations do not guarantee interoperability, they ensure that data and metadata are increasingly accessible to tools such as mediators and brokers. Through mechanisms like the Funded Projects Requirements document, barriers to interoperability across domains will be lowered over the life of the EC program.

### **Recommendation 2**

*Articulate a concise and clear definition of EarthCube that embraces the innovation aspects of the program as well as the concrete objective of creating usable infrastructure. The program definition needs to embrace discovery in both data science and geoscience. At the same time a vision and suitable metrics are needed for evaluating research products and leveraging the best results to develop enabling cyberinfrastructure.*

### **LC Response and Actions Taken**

1. **An Evolving EarthCube Definition:** The LC agreed with the RSV report that the scope of EarthCube remains unclear and ambiguous to many in spite of the existence of community-vetted mission and vision statements. The LC recognizes that the definition of EarthCube evolves as the community and the program evolve. NSF initially preferred a general and broader language so as to not predefine what EarthCube could be. Therefore the definition of EarthCube was expected to be refined over time. At its March 2016 face-to-face (F2F) meeting, the LC agreed on a more concise definition that reflects the vision as a science focused initiative and the status of the program:

***“EarthCube is a community effort to promote interdisciplinary geoscience by enabling technology, organization, and culture that facilitates connectivity through standards and protocols to existing and emerging resources.”***

2. **Definition of relevant terms:** The EarthCube definition and relevant terms including “Architecture” and “Infrastructure” as they apply to EarthCube were further defined by participants of the Architecture workshop, which was held in May 2016. The new definitions are included in the [final workshop report](#), along with the definition of System of Systems (SoS) as detailed in the response to Finding 2.3.
3. **Future Steps:** The LC will ask the Technology and Architecture Committee (TAC) to establish an EarthCube Standards Authority and develop a charter for it.

## 2.2. Finding: Transition to an LC-empowered Governance Model

### **Recommendation**

*We recommend an expedited shift in governance strategy to include a substantial top-down component. It is appropriate to consider this new model as part of “Phase II” of the program and it should be a near term goal. We recommend that the LC be empowered to make critical decisions about architecture implementation and become directly involved in defining proposal solicitations from NSF.*

### **LC Response and Actions Taken**

1. **Clarify responsibilities & authorities:** The LC worked with the NSF to clarify responsibilities and authorities in order to make progress and move forward with developing the program. This was done after the new ESSO was put in place.
2. **Stronger LC Leadership:** The LC has also taken steps to provide stronger leadership in the evolution of the EC initiative. Among these are:
  - a. the appointment of Tiger Teams, as needed, to target specific areas;
  - b. the recommendations provided to NSF for Funded Project Requirements and focus of future solicitations;
  - c. focus meetings on making decisions regarding processes and outcomes rather than information exchange.
3. **Improved LC efficiency:** Over the summer of 2016, the EarthCube Science Support Office (ESSO) was created, with a stronger mandate to lead the program. The more proactive support and leadership of the ESSO is helping the LC to increase its efficiency, ensuring interconnections among EarthCube activities, and engagement and communication with the EarthCube community for the benefit of the program. Recently, the ESSO has filled the needed support positions to realize this goal. So far three full time positions have been hired - a project manager, project coordinator, and a communication specialist.
4. **Review of EarthCube governance:** The LC held another F2F meeting on November 17-18, 2016, and spent a substantial part of the meeting on reviewing extensively the current governance structure. A document summarizing the recommendations from those discussions is currently under development. These recommendations will also address the need for guidelines to simplify the systematic- and self-assessment process of funded projects.
5. **Future Steps:** Based on the outcomes of the F2F LC meeting in November 2016, the LC will review and amend the EarthCube charter and bylaws to implement changes to the governance structure that were deemed necessary.

### **Recommendation**

*In the long term, an entirely different governance model may be needed that involves a selected set of senior stakeholders in each discipline area, who would be empowered to lead capacity building efforts within their discipline while also facilitating coordination with other discipline leaders. The UNOLS*

*program was cited as a potential organizational template for EarthCube, and there may be other models worth considering. This issue is revisited in Section 4.*

### **LC Response and Actions Taken**

1. **Future Steps:** The LC has so far focused its discussions and actions on the need for short-term adjustments to the EC governance structure that shift decision-making power in ways that improve EC's ability to meet the implementation plan. A revision of the charter and by-laws was decided at the November 2016 F2F LC meeting and that work will be completed by the end of 2016. Further discussions and strategic planning for the long-term governance model of EarthCube will take place in 2017.

### **2.3. Finding: "System of systems" is a critical concept in the EarthCube program, but it is not clearly defined.**

#### **Recommendation**

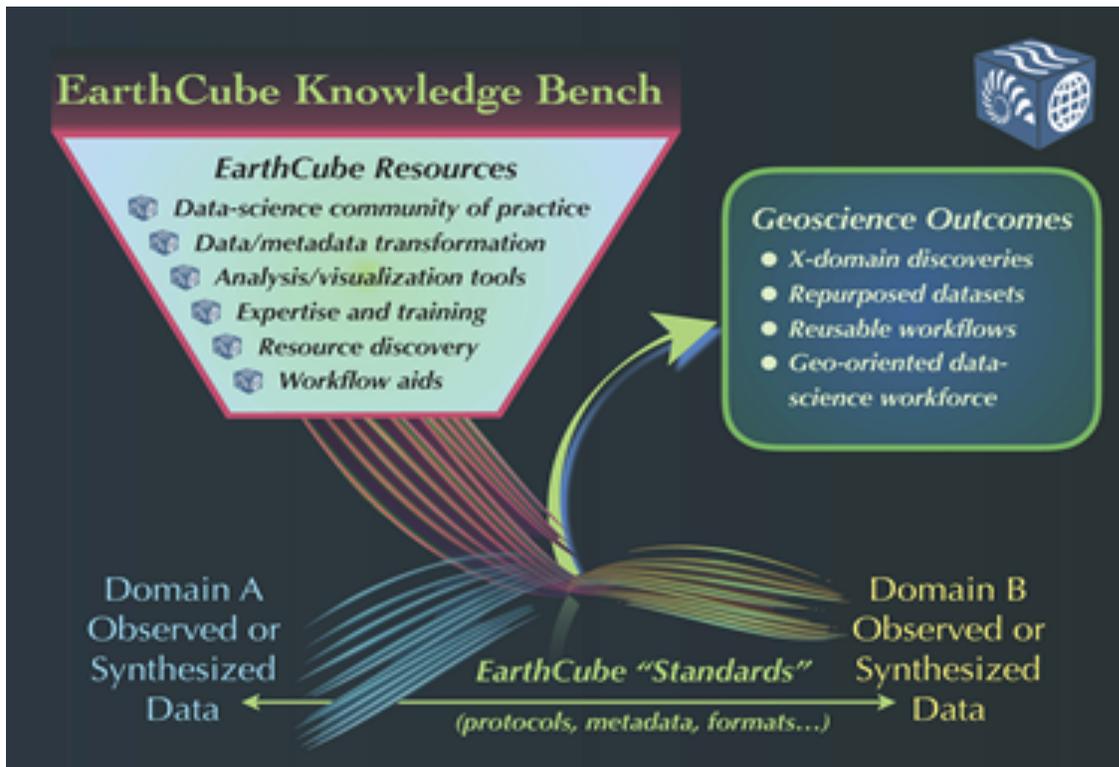
*If "System of Systems" represents the guiding paradigm for realization of the kind of data-enabled science advanced by EarthCube, then the phrase must be defined in a way that lends itself to an implementation strategy. The AC understands the phrase to refer to the aggregation of multiple heterogeneous data repositories and tool sets under a single EarthCube system. Interoperability is the cornerstone objective in this approach. The LC is free to correct this understanding if it is incomplete or in error.*

### **LC Response and Actions Taken**

1. **Architecture Workshop:** At its March 2016 F2F meeting, the LC resolved to address the need to develop an implementation plan re-allocating funds remaining in the ECTEG (EarthCube Test Enterprise Governance) budget and in budgets of different governance bodies to support a workshop to build consensus on an architecture framework and deliver a definition of the EarthCube System of Systems. The LC established an Organizing Committee under the leadership of LC member David Arctur and TAC Co-Chair Jay Pearlman, assisted by external facilitators (Matt Tricoli and Jim Barrett from Xentity). The May 2016 Architecture Workshop (AW-2016) brought together a limited number of selected EC community members, stakeholders, and outside experts in large cyberinfrastructure architectures. The final report is available [here](#).
2. **Definition of the EarthCube System of Systems:** The AW-2016 participants were charged by the LC to define the EarthCube SoS and to recommend and prioritize requirements for realizing the SoS. In the final report, the SoS was defined, with the understanding that EarthCube will operate as a collection of independently managed systems that are networked with the shared goal of facilitating interdisciplinary research. Systems will be expected to be interoperable via the use of standard interfaces endorsed by the EarthCube Standards Authority discussed earlier in this report, though data facilities that cannot easily adapt to such standards may need adapters such as brokering tools to connect.

**Definition of EarthCube as a SoS:** *“EarthCube (referred to alternatively as EarthCube or EC) is a program and a supporting System of Systems (SoS) infrastructure that enables the geosciences research community to discover and access and connect resources with a minimum of new cyberinfrastructure development. EarthCube is not intended to be a new data center, but will build on existing infrastructure and support activities to connect data, analysis and resources across the geosciences. Supporting these integration activities requires organization, partnership agreements, adoption of standards and best-practices, in addition to software resources and tools. This document outlines the basic terms, principles and the concept of operations for this EarthCube infrastructure. The infrastructure is to become a System of Systems that integrates new and existing resources and components into a framework for increasing innovation, efficiency, and effectiveness of the research enterprise. The EC SoS will promote reuse by facilitating bridging between existing components but will provide development of new components as necessary, while avoiding duplication of effort. As an operational construct, an EarthCube system of systems is “a composite system in which the component parts are heterogeneous, independently managed, and independently operable systems on their own, but are networked (loosely coupled) through EarthCube-endorsed standards to achieve shared goals.” What distinguishes a system of systems from a random collection of systems is the condition that the component systems are networked to achieve shared goals.*

3. **Architecture Implementation:** There was consensus at the AW-2016 workshop that the recommended infrastructure will be both feasible for implementation in the near term and will have a significant and enduring impact on the way that cross domain research will be done. It was felt that there was an urgent need for EarthCube to move forward with a user-focused architecture constituted as a Geoscientist’s Knowledge Workbench that can be evaluated by the geosciences community. The Workbench would provide not only tools and data access, but a collaborative environment for innovative ways to address the [Grand Challenges](#) in science research.



**2.4. Finding:** The competing “innovation” and “production” elements of EarthCube need better articulation to alleviate the ambiguity and lack of realistic program scope sensed by the Geoscience community.

**Recommendation**

*Articulate the parallel “production” and “innovation” tracks within the program more directly. Innovation is embodied in the “Building Blocks” solicitation and in the data-enabled geoscience advanced by leveraging the most effective data science products in a functioning architecture.*

**LC Response and Actions Taken**

1. **Integrate ‘Production’ and ‘Innovation’ with the Architecture Implementation Plan:** Any LC actions on this recommendation were put on hold until the architecture implementation plan became available. At its November 2016 F2F meeting, the LC decided to charge the TAC and the Council of Data Facilities (CDF) with a review of the [Solution Architecture](#) and [Architecture Implementation Plan](#) and to provide feedback that will help articulate the parallel production and innovation tracks in EarthCube.

## 2.5. Finding: Insufficient emphasis on enabling cross-disciplinary science through EarthCube.

### **Recommendation**

*Most of the defined Use Cases concern ways in which data science enables discovery within a single sub-discipline of geoscience. More emphasis should be placed on enabling truly cross-disciplinary geoscience research.*

### **LC Response and Actions Taken**

1. **Analyze cross-disciplinary Use Cases:** The RSV report expressed the view that more emphasis should be placed on enabling truly cross-disciplinary geoscience research. The LC created a charge for the already ongoing TAC Use Cases Working Group to analyze the current state of the use case collection within EarthCube, to prepare a report defining the needs and resources to collect cross-disciplinary use cases, and review the results. In September of 2016, the Use Case Working Group outlined 49 use cases, and is currently rating them according to their applicability to the cross-disciplinary charge. At the time of writing, 16 were rated as having a high level of inter-disciplinarity and a further 11 were rated medium. A significant proportion are still pending, awaiting response from the use case contributors.

## 2.6. Finding: Strategic “Imperatives” are ambiguous

### **Recommendation**

*The imperatives in the strategic plans should convey the most important things needed for the program not to fail. These should be specific, and extraneous imperatives removed.*

### **LC Response and Actions Taken**

1. **Prioritization of tasks:** At its F2F meeting in March 2016 the LC identified four aspects of the program that it considered of the highest priority for the success of the program, and has focused its actions on those aspects. These items include the development of a Solution Architecture and the Architecture Implementation Plan, identification and further development of cross-disciplinary use cases, strengthening the role of the LC and ESSO in the EarthCube governance, and ensuring better alignment and integration of funded projects into the overall development of the EarthCube System of Systems. These are the actions described in this response document. The LC will shift focus and revise the strategic plan following the comments made by AC.

### 3. ROADMAP

*Charge: What is the quality of the roadmap? Are the short, medium, and long term plans and actions appropriate for meeting the strategic vision? Are the plans achievable? How well has the roadmap, particularly the development of architecture, been informed by and incorporated existing resources, EarthCube community activities and funded projects?*

#### 3.1. Finding: The program lacks a concrete 5-year implementation plan.

##### **Recommendation 1**

*The Roadmap documents should be supplemented, or converted to, an “implementation plan” or “implementation roadmap”. This document should include specific prioritized milestones, and clearly defined metrics against which program progress is to be measured.*

#### **LC Response and Actions Taken**

1. **Development of the Architecture Implementation Plan:** The development of an architecture for EarthCube and a corresponding implementation plan was considered as the highest priority by the LC. To start the process, the LC charged an Organizing Committee under the leadership of LC member David Arctur and TAC Co-Chair Jay Pearlman with the help of external facilitators (Matt Tricomi and Jim Barrett from Xentity) to organize the Architecture Workshop (AW-2016) that brought together a limited number of EC community members, stakeholders, and outside experts in large CI (Continuous Integration) architectures. The final report (available [here](#)) recommended an architecture framework that was considered both feasible for implementation in the near term and to have a significant and enduring impact on the way that cross-domain research could be done in the future. It was felt that there was an urgent need for EarthCube to move forward with a user-focused architecture constituted as a **Geoscientist’s Knowledge Workbench** that can be evaluated by the geosciences community. The Knowledge Workbench would provide not only tools and data access, but a collaborative environment for innovative ways to address the [Grand Challenges](#) in science research.

The ESSO was then charged by NSF to create an Architecture Implementation Plan based on the high-level architecture defined in the Architecture Workshop [Report](#) (See Section 2.3 above). The LC established an Architecture Implementation Plan Tiger Team (AIP-TT) that was [charged](#) to work with an external contractor selected by ESSO ([Xentity](#)). Xentity had facilitated the AW-2016 and agreed to quickly develop a plan to move the recommended architecture framework towards an actionable Implementation Plan. David Arctur, who was the LC member who co-led AW-2016, was appointed by the LC to chair the AIP-TT, which began regular meetings in September, 2016. The LC also appointed to the AIP-TT representatives from each of the standard committees (TAC/SC) and the CDF.

Based on the AW-2016 outcomes and considerable input from the EarthCube broader

community through the development of a [reference library](#)<sup>1</sup>, and with guidance and review provided by the AIP-TT, Xentity developed a more detailed [Solution Architecture](#), which was delivered to ESSO in October, 2016. An [Architecture Implementation Plan](#) based upon the adoption of the Solution Architecture was delivered to ESSO by Xentity on November 17, 2016. The AIP-TT prepared [discussion topics](#) relating to the two documents. The LC is currently preparing to solicit feedback on the Solution Architecture from the TAC and the CDF. This feedback from the community, including the newly formed EarthCube Advisory Group, will be sought to form a basis for the LC to assess both the Solution Architecture and the Architecture Implementation Plan and develop an actionable 5-year plan by March 2017.

The Draft Implementation Plan identifies four phases for completion, prior to the release of EarthCube v1.0.

- **Project Phase 0 – Architecture selection** - The first phase is architecture selection, and should follow a project management model. This is a phase, as the winning pattern prototype work should move immediately into Phase 1 and should be done with Phase 1 scope in mind
- **Project Phase 1 – Initial Operating Capability** - Should have multiple sprints using Agile Project Management with the Architecture Runway providing input to get to the first key epics developed, decomposed into user stories, and tightly managed in short quick sprints to demonstrate rapid results.
- **Project Phase 2 – Continuous Improvement – Supplier Integration Driven** – This phase should focus on plugging in the prioritized system interfaces for data sources in and data interfaces out. Though some of this will be in phase 1, it is critical at this stage to engage the other suppliers to assure the progression is made in lockstep with each other, and not diverging from the architecture and principles due to delays, political pressures, or possibly, technical exuberance.
- **Project Phase 3 – Continuous Improvement – User Feature Driven** – At this phase, the project should now be moved into being driven less by project and architecture initial framing and more by community features that are needed. This phase essentially marks the version 1.0 and future phases should be minimized now into sprints as the project should move into operational continuous improvement mode.

The Architecture Implementation Plan includes a specific set of recommended milestones (Section II.2.1) proposed for the program. The LC at its November 2016 F2F meeting developed a process to seek community feedback on the Architecture Implementation Plan and incorporate this feedback. This feedback process requires a rapid turnaround. Feedback is being directly solicited from the TAC and CDF Committees, and simultaneously from the LC. Responses from the TAC and CDF will be synthesized by the LC for feedback to Xentity (if necessary), or for adoption by the LC for EarthCube Governance.

---

<sup>1</sup> A summary of reference material provided by the EarthCube community to Xentity can be found [here](#).

### **Recommendation 2**

*The documents should also make the period of consideration clear. In this case, the implementation roadmap is understood to be a 5-year plan. It has been said that the “path to infrastructure takes decades,” but we need a more concrete implementation plan to justify continuing the program beyond the next 5 years. Short-term explicitly defined solutions are needed to avoid volunteer fatigue and external criticism.*

### **LC Response and Actions Taken**

1. **Timeline provided in the AIP:** The [Architecture Implementation Plan](#), developed by Xentity and currently under review by the TAC, CDF and the LC, provides a concrete timeline to 2020 described in Section II.2. of that document. The Timeline sets a deadline of 9/3/2017 for implementation of Phase 0, the Architecture Implementation, 8/29/2018 for Initial Operating Capability (Phase 1) of the EarthCube Architecture, with the first phase of “supplier” integration (inclusion and integration of funded projects) ending 2/25/2019, with Continuous Improvement (Phase 2) continuing throughout. Phase 2 completion is expected by 5/26/2019, and Phase 3 completion expected at 2/20/2020.

### **Recommendation 3**

*While tackling the above imperatives, the EC needs to retain some of the grassroots aspects necessary to allow participation of communities that are just beginning to try and understand how cyber-infrastructure can assist in the development of databases, new knowledge, or new capabilities to extract knowledge.*

### **LC Response and Actions Taken**

1. **Enumeration of EC Standards:** The enumeration of EC Standards through the [Standards Compilation spreadsheet](#) and the compilation of existing Use Cases provides an opportunity for the transparency required to facilitate “grassroot” community engagement. The [Software Assessment Spreadsheet](#) provides a clear guideline for operational requirements as new communities begin the onboarding process, or commit to EC through the process of writing grants. This spreadsheet is a synopsis of the larger [ESIP Software Guidelines](#), (Scott et al., 2016), which in turn is based on the [Software Evaluation Guide](#) (Jackson et al., 2011) developed by the Software Sustainability Institute. The importance of Use Cases in driving Architecture Implementation and in Gap Analysis for Architecture and Workbench development (Architecture Report Section I.1.5). Currently the Architecture Report provides a mechanism to use feedback from the Workbench to help steer future solicitations. In this way, the development of the Geoscientist’s Knowledge Workbench, its documentation and its practical implementation will provide a clear impetus for domain user groups to engage with EarthCube and develop the EarthCube architecture.

### 3.2. Finding: The development of functional infrastructure is impeded by unrealistically broad scope, lack of standards, and lack of metrics.

#### **Recommendation 1**

*There is a need to define and/or adopt specific standards that ensure projects are integrable and data repositories are interoperable. These standards should be established by a small group of technologists and geoscientists, informed through discussions with the broader community. All grant awardees must be required to follow these standards.*

#### **LC Response and Actions Taken**

1. **EarthCube Standards Authority:** In 2015, the EarthCube Technology and Architecture Committee (TAC) had set up a Standards Working Group (WG) to address specification of and adoption of technology standards for metadata, software and interoperability for the program. The Standards WG released a [report](#) with recommendations in October 2015. Following the RSV, the Leadership Council asked the Standards WG to generate an implementation plan for technology standards for funded projects and organize a session at the All-Hands Meeting 2016 to discuss this plan. The Working Group proposed that EarthCube should not be a formal standards development organization, but should instead have a permanent “standards authority” for the purposes of finding, recognizing, and promoting use of relevant and useful standards and practices. The EarthCube Standards Authority, also recommended in the Architecture Implementation Plan, would work with EarthCube funded projects to facilitate the adoption of standards and systematic practices, which will provide mechanisms for integrative activities.

The work of the EarthCube Standards Authority would also inform and leverage the architectural development of EarthCube, as appropriate. Following on from the AHM, the working group made several recommendations:

- EarthCube needs a registry for specifications of EarthCube practice; the barrier for entry into the registry should be kept low—focused on the quality and completeness of documentation and testing.
- Utilize a tiered scheme to categorize specifications in different levels of adoption/recommendation
- Collect feedback and evaluation on the utility, usage, and performance of registered specifications and use this to categorize adoption/recommendation levels; this must be a transparent process.
- Implement validation processes for recommendations with a high level of adoption.
- EarthCube governance should constitute a permanent “standards authority” to carry out the evaluations and recommendations of the Standards WG.
- The hierarchical and versioned nature of standards will require development of a structured taxonomy allowing specific customizations and profiles to carry the context of the parent standard(s).

In addressing the initial recommendations, [a live shared spreadsheet](#) was created to act as an interim registry for standards for EarthCube.

2. **Assessment Framework:** The LC supported activities to design an Assessment Framework for the various components of EarthCube, including funded projects, governance, community engagement, and overall impact. The software assessment framework addresses the EarthCube Strategic Vision by promoting: best practices in software development, complete and useful documentation, interoperability, standards adherence, open science, and education and training opportunities for research developers. It will also provide the LC and NSF with guidance on resources needed to implement the Assessment Framework. This process was started with a joint EarthCube/Federation of Earth Science Information Partners (ESIP) Workshop held June 1-3, 2016 to define an assessment framework for software, The final workshop report assessment framework is posted online (Software Assessment Spreadsheet). This was the first step in developing the assessment framework.

Further assessment mechanisms are described in the proposed **Solution Architecture** and incorporated into the design of the Geoscientist Workbench. These mechanisms will assess resource interoperability, resource use, user feedback and resource maturity and provide results in a report available through the workbench. These assessment results can be compiled to provide summary metrics for the Workbench as a whole or components therein.

### 3. Future Steps:

- a. Implement the Software Assessment tools in the formal evaluation of project outcomes and annual reviews.
- b. Development of milestones and metrics

### **Recommendation 2**

*Based on these standards, there will be a need for solicitations to promote implementation of the “system of systems” approach. The LC should look to “low-hanging fruit” in this effort—for example, areas where quality data are not utilized, or obvious synergies are not exploited.*

### **LC Response and Actions Taken**

1. **Funded Projects Requirements:** Part of redefining the strategic vision for EarthCube is determining which projects best fit the new program mandate. Rewriting project award criteria has been advanced by several different initiatives during the response period. The LC formed the Funded Projects Requirements Tiger Team (FPR-TT), led by Yolanda Gil and David Arctur, charged to develop guidelines for newly funded EarthCube projects so they would better integrate with EarthCube activities. The FPR-TT took into consideration the report from the Architecture Workshop, which covered architecture aspects in terms of high and mid-level requirements, as well as guidance for future solicitations and new awards. It also considered the report from the Assessment Workshop described below.

The FPR-TT generated two reports that were approved by the LC and submitted to NSF. The first one released in June 2016 addressed the funding cycle for 2016. The second report provided recommendations for project requirements for future year awards. Both included a

set of 15 requirements for new project proposals to fulfill with the goal to better align future work with the scope and definitions created for EarthCube. They were designed to promote best practices, therefore enabling integration in the EarthCube system of project products, supporting coordination across project activities, and facilitating assessment and dissemination to the user community. Given that the EarthCube architecture, standards, and other key aspects of those requirements are still underspecified, the reports suggest that appendices be created to specify in more detail how the requirements can be met as EarthCube evolves. Feedback from the community was collected for both reports through EarthCube governance and at the All Hands Meeting in July 2016.

- Interim report with [Guidance for 2016 Awards](#)
- Final Report with [Guidance for Future Awards](#)

2. **Input to 2017 NSF EarthCube Solicitation:** As part of the process to realign project goals, the RSV Report featured heavily on the recommendations from the LC to the NSF regarding writing the 2017 EarthCube Solicitation. In particular, they targeted cross-disciplinary science, data science, and bridging the gap between building blocks and functional architecture. The specific recommendations made by the RSV report were:

- a. The 2017 Earthcube solicitation should be written to attract a broad base of innovative thinkers that cross the disciplines of data science and geoscience.
- b. The 2017 Earthcube solicitation should embrace the innovation aspects of the program as well as the concrete objective of creating usable infrastructure.
- c. The 2017 Earthcube solicitation should encourage proposals that foster innovative exploration around the explosive field of data science, while simultaneously considering ways to leverage the most applicable results for advancing geoscience research.

In September 2016, the LC provided concrete recommendations to the NSF for the 2017 EarthCube solicitation:

- A systematic approach to the assessment and a pathway towards integration of funded projects. The LC recommended a specific call within the solicitation for “developing a way to enable projects to become part of a system of systems”.
- Initial pilot/prototype projects that will work toward implementing the architecture that is expected to emerge soon with the Architecture Implementation Plan, with an emphasis on promoting advances in cross-disciplinary geoscience.

At the time of writing this response, the NSF solicitation has not been made released yet.

### 3. Future Steps

- a. The recommendations for future EarthCube solicitations include requesting proposals to address important aspects of the architecture implementation plan, and to develop approaches to integrate projects in the SoS framework.
- b. The LC will continue to update the general Funded Projects Requirements, adding

specifics as the EarthCube architecture and standards become more concrete.

- c. The LC will work with the NSF to establish a process for the LC to provide guidance to future NSF EarthCube solicitations.

## 4. COMMUNITY ENGAGEMENT

*Charge: How well integrated and coordinated is the scientific community in the roadmap and other EarthCube activities? Are the plans to engage various communities appropriate?*

### 4.1. Finding: Insufficient emphasis on data science objectives in the roadmap

#### LC Response and Actions Taken

1. See response to Finding 2.1 on page 3 of this document:
  - a. **Emphasize EarthCube’s role in Data Science**
  - b. **Encourage Data Science in future solicitations**
  - c. **Interoperability as a Funded Project Requirement**
2. **Architecture Implementation Plan:** The implementation of the Geoscientist’s Workbench will provide an exploratory environment for understanding the utility of existing technologies and identifying opportunities for new research in informatics and data science. The opportunity to use EC provided resources (e.g. cloud environments, HPC, large volume datasets) in an exploratory environment has great potential for advancing data science in ways we can’t imagine.

### 4.2. Finding: The gap between the Building Blocks projects and a functional “EarthCube Architecture” is enormous.

#### **Recommendation**

*This gap is likely not bridgeable through a community governance model, as discussed previously. A hand-picked group of experts is likely needed to make hard but informed choices based on results from funded projects. The two-year duration of the awards was also deemed unlikely to lead to successful implementation in an architecture. The AC commented that the development of an implementation roadmap may be premature. The LC may wish to develop this document after the products from the Building Block projects can be evaluated later in 2016. This exercise will determine which products are most likely to move the EarthCube architecture forward.*

#### LC Response and Actions Taken

1. **Development of the Solution Architecture:** The development of the Solution Architecture is expected to provide the framework and assessment mechanisms to more clearly identify the relationships between Building Blocks and a functioning infrastructure (i.e. the Geoscientists Workbench). Assessment mechanisms are described in the proposed Solution Architecture

and incorporated into the design of the Geoscientist Workbench. These mechanisms will assess resource interoperability, resource use, user feedback and resource maturity, and provide results in a report format made available through the workbench. The assessment results can be compiled to provide summary metrics for the Workbench as a whole and components therein including funded projects. The intention is not to preclude contribution to the workbench capabilities outright, but rather provide a test environment that allows contributed capabilities to demonstrate their own value as determined by users. This approach will require that capabilities be interoperable with the basic architecture, and therefore the LC recommends that this role be mediated by technical development staff provided by ESSO. In addition, an ESSO Technical Officer will oversee the engagement between the capability providers (e.g. building blocks, data facilities, third-party organizations) and the ESSO software development team.

The level of interoperability of individual capabilities will be determined by the adherence to EC recommended standards (Finding 3.2) as well as the quality of the software/dataset/metadata. The ESSO technical team will work directly with capability providers to facilitate entry into the workbench environment at the lowest cost of entry to the provider. The ongoing value of the capability will be determined by use in the workbench. Value may also be determined by ongoing improvement and evolution of the capability.

Integration of funded projects into the Solution Architecture is additionally addressed within the [Solution Architecture](#) report (discussed in Section 3.1, above), in particular from Section 1.1.4.3: “**EarthCube Development Portfolio Investment Candidates for Re-use:** *“Governance would need to consider which BB and IAs could be candidates as a workbench capability, part of the EC Framework services, or provide resource feed to partner aggregator.”*”

A chart is also included in Section 1.1.4.3. that conceptually outlines existing funded-project roles in the defined architecture.

2. **Assessment Framework:** The LC supported activities to design an Assessment Framework for the various components of EarthCube, including funded projects, governance, community engagement, and overall impact. It will also provide the LC and NSF with guidance on resources needed to implement the Assessment Framework. This process was started with a joint EC/ESIP Workshop to define an assessment framework for software quality, held June 1-3. The [final workshop report](#) assessment framework is posted online ([Software Assessment Spreadsheet](#)). This was the first step in developing the assessment framework.

### **4.3. Finding: Proposal solicitation structure may not be conducive to broad interdisciplinary (GEO-CISE) involvement.**

#### **Recommendation 1**

*A different approach to the solicitations involving feedback and oversight from the LC could lead to substantially increased community involvement at the same funding level.*

## LC Response and Actions Taken

1. **Funded Projects Requirements:** see response to Finding 3.2, Recommendation 2, on page 14
2. **Input to 2017 NSF EarthCube Solicitation:** see response to Finding 3.2, Recommendation 2, on page 15.
3. **Future Steps**
  - a. The LC will continue to update the general Funded Projects Requirements, adding specifics as the EarthCube architecture and standards become more concrete.
  - b. The LC will work with the NSF to establish a process for the LC to provide continued guidance to future NSF EarthCube solicitations.

### *Recommendation 2*

*Funding a smaller dedicated group within the community may lead to the concrete infrastructure products that will ultimately make the case for a much broader group of scientists to engage in EarthCube.*

## LC Response and Actions Taken

1. **Solution Architecture:** As part of the development of the Solution Architecture, the LC recommends that software development staff should be provided by ESSO (see response to AC Finding 4.2 on page 17). The ESSO Technical Officer would oversee the engagement between the capability providers (e.g. building blocks, data facilities, third-party organizations) and the ESSO software development team.

## 4.4. General Recommendations

### *Recommendation for LC*

*Consider whether there is an opportunity or sufficient knowledge to be gained by examining some of the end-user workshops that reported out well – to extract best value or best practices out of the best end-user-reports*

### **LC Response:**

Beginning in summer 2012, NSF funded a series of 24 EarthCube domain end-user workshops engaging some 1650 participants. Template-structured reports from each of these workshops were reviewed by parts of the EarthCube demonstration governance twice in 2015 (once by the Science Committee (SC) as part of the scientific visioning exercise that culminated in the “2020” document, and once by the TAC, as part of the Use Case collection effort). The SC effort resulted in significant findings about common themes, cyberinfrastructure needs, and data availability issues (as summarized in the 2020 document). The LC agrees with the implication of the RSV recommendation that some of these reports indicate higher levels of sophistication and awareness of the types of activities and platforms that might be provided to them by the EarthCube effort, and emergent

practices with data handling and documentation that could be shared with other less adept domains. It also agrees that more could be done with these reports in aggregate to help direct the current EC funded projects community and NSF program managers towards high impact activities that would serve end-users well.

**Future Steps:** The LC in conjunction with the SC (and with input from the TAC and cognizant NSF PMs in the relevant programs) plans to initiate another data mining activity on these reports, to be followed by a coordinated effort to re-engage with a subset of these end-user communities (via representatives). We will focus on groups that are most poised to take advantage of emerging EarthCube capabilities to support high impact science. We envision a workshop activity coupled with a planning effort, perhaps aided by a Tiger Team, to both gather end-user community input via representatives and to identify scientists that are ready and willing to engage directly with either currently funded project, or new projects via an enhanced EC solicitation.

### ***Recommendation for NSF***

*Future solicitations should contain input from the leadership Council about the standards and formats that must be considered when producing or integrating data sets that will foster community knowledge of the constructs needed*

## **LC Response and Actions Taken**

1. **Funded Projects Requirements:** The LC was invited by NSF to provide guidelines and requirements for Funded projects (see response to established a Tiger Team that was charged to generate these guidelines for projects funded in 2016 and for future funded projects).
2. **Future Steps:**
  - a. The LC will continue to update the general Funded Projects Requirements, adding specifics as the EarthCube architecture and standards become more concrete.
  - b. The LC will work with the NSF to establish a process for the LC to provide guidance to future NSF EarthCube solicitations.

## **5. NEXT STEPS**

*Charge: What next steps should be taken by the EarthCube community? What are gaps or missing elements of the strategic vision and roadmap? Are there other communities or initiatives the EarthCube community should consider?*

### ***Recommendation***

*Write down the guiding principles for what is clearly meant by the system of systems and what standards it would need to address so that the community expectations can be channeled. Then get*

*down into prioritized items that address both sides of the problem, the innovation gathering of science communities into exploiting the data assets (analysis world) as well as influencing the NSF awards solicitations where it is appropriate.*

## **LC Response and Actions Taken**

The LC undertook a series of 'Next Steps' that have been highlighted above in this document and align with the recommendations of the AC. These include:

1. **Defined the System of Systems:** see response to Finding 2.3 on page 6. This definition was provided by the Architecture workshop that was organized
2. **Produced the Architecture Implementation Plan:** see response to Finding 3.1, page 10. The AIP is now under review by the TAC and CDF.
4. **Recommendations to NSF for 2017 solicitation:** see response to Finding 3.2, on page 15.
5. **Next steps toward the Solutions Architecture and Implementation Plan:** The Solutions Architecture and separate Implementation Plan document prepared by Xentity have been discussed in previous sections. These two documents lay out general guidelines to help EarthCube move forward and provide a helpful starting point for next steps. These documents consider many further necessary next steps as being out of scope, primarily because they require additional work by the EarthCube community and its governance. The LC will consider a document generated by the AIP-TT to decide in further next steps.

## ***Recommendation for the LC***

*It appeared to the panel that the LC did not up to this point have a viable structure or mandate to produce an implementation plan. This needs to be corrected. The panel challenges the Leadership Committee to solve this dilemma, perhaps through a more authoritative management construct in the Project Support Office under the guidance/direction of the LC.*

## **LC Response and Actions Taken**

See responses to Findings 2.2 on page 5:

1. **Clarify responsibilities & authorities**
2. **Stronger LC Leadership**
3. **Improved LC efficiency**
4. **Review of EarthCube governance**
5. **Future Steps:** Based on the outcomes of the LC F2F meeting in November 2016, the LC will review and modify the EarthCube charter and bylaws to implement changes to the governance structure that were deemed necessary.

### ***Recommendation for LC***

*The panel concluded that the leadership council should be prepared to meet the panel again in about 9 months to report on progress.*

### **LC Response and Actions Taken**

The LC is preparing to meet with the newly formed Advisory Group in early 2017, following the AG kickoff meeting in late November.

### ***Recommendation for LC and the NSF***

*In the end, the AC was pessimistic about the ability of EarthCube program to evolve into a thriving enterprise in its second five years under the current organizational model. What is likely needed in “Phase II” is a new structure, deliberately designed to realize the requested implementation plan. This structure should involve a select set of experienced community leaders from each of the GEO discipline areas. These leaders would be empowered to build data science capacity within their disciplines, while also coordinating with other leaders to work toward realizing the (still undefined) EarthCube vision. The detailed implementation of this new governance model should await a broader review of the EarthCube Program, as discussed in Section I.*

### **LC Response and Actions Taken**

As shown in the earlier sections, over the past several months, EarthCube has taken definitive steps towards better defining near- and long-term objectives and scope. These steps, facilitated by an Office that has been only recently staffed, enabled progress towards a specific, detailed plan for implementing an infrastructure within the next five years. The LC is confident that EarthCube could begin implementation by May, 2017. This infrastructure provides both an operational infrastructure to enhance discovery, brokering and access to resources and a workspace to enable innovation by providing an integrative approach across communities of resource providers and consumers.

### **Future Steps**

- The LC needs to adapt, adopt and promote the Implementation Plan and develop the governance requirements to support the implementation. This should begin early 2017.
- The recommendations for future EarthCube solicitations discussed above include requesting proposals to address important aspects of the architecture implementation plan and to develop approaches to integrate projects into the SoS framework.
- Develop milestones and metrics in the implementation process and specify the roles and requirements in enforcing security measures for the solution architecture.
- Establish an EarthCube Standards Authority
- The initial Architecture Report recommended that “The first phase of the architecture could be operational in about a year from initial funding and include interfaces with EarthCube CDF data facilities and use cases defined by scientists.” More effort should be devoted to

developing the use cases and drivers for the architecture, not only for its first phase but in future phases.

- Implement the Software Assessment tools in the formal evaluation of project outcomes and annual reviews.
- The LC is preparing to meet with the newly formed Advisory Group in early 2017. The Advisory Group was established to move towards oversight and guidance from “a select set of experienced community leaders from each of the GEO discipline areas”.

***Observation:** The panel also noted that the available documents were silent on security. Activities that are intended to integrate across boundaries or federate across institutions and geographies need to be aware of and conscious of the ever-increasing need to align with the Federal Information Systems Management Act (FISMA). FISMA is best postured on the front end of a development with an ultimate objective of withstanding third party audits so that EarthCube complies with “best practices” security guidelines.*

#### **Future Steps:**

- The EarthCube architecture development has just started now. During implementation of the architecture and infrastructure, security will be given high consideration. A first step has already been an action by the ESSO to address the lack of secure access to the EarthCube website. This was completed two months ago.

## **6. PROGRAM SUSTAINABILITY**

### **6.1. Finding: It is difficult to assess at the present juncture whether EarthCube will be sustainable.**

#### **Recommendation**

*The recommendations outlined above involve a transition to Phase II. They include the need for succinct definition of EarthCube, changes in the LC’s governance approach to structured rather than consensus driven decision making, restructuring the process for more focused definition of program solicitations, and producing a bona fide implementation roadmap. These are seen as prerequisites to adopt best practices, communicate standards, and evolve to a production track. Only then can a sustainability plan be defined.*

*The AC feels that it is essential for the LC to move rapidly into Phase II as defined by the above actions. A useful timeline for responding to the AC recommendations is by the time of the All Hands Meeting in July so that a community communication plan can be implemented. The next AC meeting should involve feedback on progress and should be held in September this year.*

#### **LC Response and Actions Taken**

Sustainability of EarthCube depends on how effectively it becomes a solution for geoscientists in

discovery, access, to and usability of geoscience data sets and models. Execution of science-driven workflows within an EarthCube framework is also essential as this will enable code reuse and improve efficiency of geoscientists in doing science and not developing new workflows to address existing problems. The EarthCube SoS will enable geoscientists to improve their efficiencies, reducing the amount of time needed to assemble and transform the data into usable formats and into a processing environment that is ready for conducting their science. Making EarthCube sustainable is to make it relevant and a useful resource to the broad geoscience community, something it has not yet demonstrated.

As described above, the design of an EarthCube architecture is currently under review by a broad community including the EarthCube TAC and the CDF. At the November 2016 F2F LC meeting there was strong support for taking specific actions to move EarthCube into an implementation phase as recommended in the RSV report. The implementation phase will be when many of the EarthCube developed building blocks and conceptual designs can begin to be merged with the emerging architecture. The architecture has a key component called the workbench that will result in much of the new enabling architecture EarthCube needs to help geoscientists.

1. **Soliciting feedback from the CDF:** Since many of the resources that EarthCube will draw from those existing within the CDF, the LC is soliciting input from the CDF to determine which workbench components they need in order for connections between their CDF organization and the EC workbench.
2. **Prioritize workbench development:** The LC will establish milestones and priorities to identify components of the workbench to be developed and the order of development.
3. **Technical Support at ESSO:** The LC agreed at its November F2F meeting that ESSO should develop a proposal to NSF to fund a few key technical positions needed to begin development of key parts of the architecture with an emphasis on the workbench. This will include:
  - a. a Chief Technical Officer, this position will enable a more structured approach to EarthCube development as well as ensuring strong connections with assets managed by the existing CDF members.
  - b. a small team of technical developers
4. **Science Liaison at ESSO:** The LC also recommended that a Science Liaison position be added at ESSO to work with EarthCube Stakeholders, particularly geoscientists, to inform what EarthCube is developing to support their science and also to identify use cases that the EC architecture will support

The longer-term sustainability of EarthCube will be realized when it has demonstrated that it improves the efficiencies of geoscientists as they pursue their research. EarthCube has not reached that point yet. EarthCube can draw from the experience of many of the CDF members who have gained a certain level of sustainability as it moves forward. The LC also feels that longer-term sustainability of EarthCube can also be improved if the technical development of the architecture and workbench are augmented with the development of a standard suite of interdisciplinary products that draw from the long tail of science data sets as well as the holdings of the CDF. These products would enable routine workflows that can produce products of general interest to the public, to educators as well as to geoscientists.

In summary, sustainability of EarthCube should not be measured by the generation of external funding but rather in how it demonstrates improved efficiencies and increased productivity for doing research in the geosciences.