

EarthCube Action Plan 2017

Version March 29, 2017

Drafted by:

EarthCube Leadership Council

Introduction

This document lays out the EarthCube Leadership Council's (LC) recommended path towards moving EarthCube from the process of defining community needs and a governance structure to implementation of an integrative cyberinfrastructure in support of the goals set by the EarthCube community and the National Science Foundation. This Action Plan is based on the outcomes of recent activities initiated partly in response to the EarthCube Reverse Site Visit of February 2016, and includes the Architecture Workshop in May 2016, the Solution Architecture and its associated Architecture Implementation Plan (AIP) provided by Xentity in Fall 2016, and the corresponding EarthCube community feedback. The Plan includes mechanisms for continued and directed development of an EarthCube cyberinfrastructure, a metrics and assessment process, improved community engagement, and the development of a long-term sustainability plan for EarthCube.

Summary of Recommended Actions

The LC recognizes the need for full-time, high level support to move the ESSO from a Science Support Office to an office that will implement an integrative, cross-domain architecture and has the ability to create and maintain authoritative connections to the resources required in support of the defined system of systems approach. The addition of a Technical Officer (TO) is being taken as the first step. The TO will have the responsibility to further develop the AIP within the first year by providing more specific, implementable recommendations towards an AIP-2.

Architecture Implementation: Phased Approach (see Priority Activities 1, 2, & 3)

- Informed by prior feedback, develop an actionable Architecture Plan that includes an EarthCube Standards Adoption process (Technical Officer - TO)
- Craft a near-term plan to address Priority Area 1 (TO)
- Release a solicitation for proposals (sub-awards with a total funding of \$300K) to build Priority Area 1 Architecture components (ESSO)
 - Identify priority activities for sub-award proposals and generate an evaluation rubric (LC)
- Engage community partners (public/private/academic) to explore available technologies and CI components to advance the Architecture Plan (LC/ESSO)

Expectations and Metrics of Success and Readiness

- Synthesize, analyze and evaluate projects and outputs, their performance, adoption, and effectiveness (LC through a Tiger Team - TT)
- Refine EC goals and expectations to better foster, promote and track the program progress towards achievable goals
 - Define EarthCube program metrics and assess progress (LC via TT)
 - Define criteria/metrics to evaluate LC effectiveness as a governing body

Community Stewardship

- Establish resource requirements and processes to help EC projects achieve goals (LC/ESSO)
 - Establish systems to ensure better on-ramp for new projects and integration with the EC community (LC/ESSO)
- Provide input to NSF for the 2017 Solicitation by July 1, 2017 (LC)
- Review and refine governance charters by AHM-2017 (LC)

In addition to these critical actions, the LC has drafted a separate document, “*Enacting the Action Plan: LC Recommendations for Resources Needed*” (see separate document) that specifies critical resources needed to move forward over the next 12 months, including the transition of ESSO from a Support Office to a Project Office, which includes the addition of an Executive Director and a Central Development Team.

Architecture Implementation:

A Parallel and Phased Approach to Address Immediate Needs

A survey of the EarthCube Community and CDF members resulted in feedback on the AIP. From this feedback, the LC has established a set of prioritized activities for the development of technical components with associated timelines. As part of the survey, respondents were asked to rank the importance of a variety of potential EarthCube services and infrastructure. Identified capabilities were grouped into three categories, which form the basis of the three Priority Activities, below. Implementation of some activities identified below could be phased as components are developed and implemented, while others, such as Resource Assessment Services and Community Networking and Training Services, must advance in parallel with the overall infrastructure development.

Priority Activity 1: Discovering and Accessing Geoscience Resources

A number of projects have been funded or directly enhanced through the EarthCube program. Projects must be discoverable, accessible, and usable over the long term, if we wish to achieve sustainability goals. Resources outside the EC sphere must also be identified. The following elements defined below have been identified as functions that can and should be achieved within the first year of the adoption of this EarthCube Action Plan.

Resource Discovery Services: Discovery of resources available to geoscientists is the keystone of EarthCube success, including access to existing and new services that may reside within EarthCube (EC) or outside via partner discovery services. Combined priority score - 6.7/8

Resources Registry: All key resources must be registered to be able to be accessible through the Resource Discovery Services. The EarthCube Architecture will need to provide this capability. Combined priority score - 6.5/8

Resource Distribution and Access Services: As data services and software become accessible, EC must facilitate access to and distribution of resources through standards-based services, including links to existing CDF resources. Combined priority score - 6.4/8

Priority Activity 2: Advancing Usability and Interoperability of Resources -- Mediation

Transformation Services: Transformation services will enable data providers and consumers to select from a variety of data formats or accepted standards-based frameworks for data distribution and access (e.g., SWE-> netCDF or visa-versa). Transformation is typically included as a service in a Brokering framework. Other transformation services can be translation of units; geo-spatial sub-setting or gridding tools, etc. Combined priority score - 4.7/8

Brokering Services: Brokering Services including mediation are implemented in a Brokering Framework that maps discipline specific data attributes (metadata, standards and data formats) from one discipline to another. This mapping eliminates the need for scientists to learn the details of data instantiation across disciplines. These services include standards and metadata transformations. Combined priority score - 4.6/8

Semantic Services: The W3C Semantic Web (RDF, OWL, LOD) implementation facilitates discovery and understanding. As the EC community develops domain-specific and cross-domain registries of terms and associated ontologies, the EC Infrastructure must provide the necessary services to create, manage and utilize the resources. Combined priority score - 4.6/8

Linking or Interoperability Services: Services and resources need to be integrated through the use of standard APIs and standardized metadata. With sufficient metadata, automatic mediation for interoperability is possible. Combined priority score - 4.4/8

Priority Activity 3: Combining Resources in Scientific Workflows

Workflow Services: Workflow services provide understanding in geoscience research, as well as reproducibility. As Workflows are registered, standards-based workflows enable the reuse of well-documented workflows with input from the broader community via Resource Assessment Services. Combined priority score - 4.0/8

Fusion and Visualization Services: These services will be a key but final component of the EarthCube Workbench. The Fusion and Visualization Services will provide a user-friendly interface for the utilization of the resources offered through the Discovery, Distribution, and Access Services developed as Priority Activities 1. Combined priority score - 3.2/8

Throughout the development of the services mentioned above, the following associated services will need to be implemented:

Resource Assessment Services: Services that enable interactive, constructive assessment of tools and data provide a significant contribution to the broader geo-science community.

Community Networking and Training Services (CNTS): Some examples of community networking and training services include instantiating single-sign-on across partner organizations and formalizing training for creation/utilization of community resources for EC registered members.

EarthCube AIP Timeline for the Next Year

ESSO is in the process of hiring a Technical Officer (TO) who will be engaged by the third quarter of the ESSO fiscal year 2017 (Mar-Apr). The first priority for the TO will be to take the Architecture Implementation Plan (AIP) generated by Xentity [\[link\]](#), consider the feedback from the LC, TAC and CDF, and generate a more specific, actionable Architecture Implementation Plan (AIP-2).

The TO will provide a proposal for addressing Priority Activity One within 6 months of hiring. This proposal will be informed by existing Architecture reports and workshops, with input from LC, TAC, CDF and other committees, but driven by the TO. The proposal will be vetted by the Leadership Council - without further feedback from the EC committees - and the TO will then be empowered to move forward with the implementation of solutions for Priority Activity One, with support from a small (\$300K) award to ESSO, to be distributed as sub-awards for the components needed to implement the defined services.

Once the AIP has been further refined, the LC will approach community partners, identified in part through the [Liaison Team's Mapping the Landscape](#) effort, in an effort to leverage existing community resources and capabilities, identified through the TO's Architecture Plan. These partnerships will be formed as an effort to develop solutions for long-term sustainability and will leverage existing resources from external organizations. Crafting the partnerships will be supported through administrative assistance provided by ESSO.

EC Architecture Refinement Workshop: This activity is being planned to bring together funded projects (past/present) and CDF members to work with the soon-to-be appointed TO in further refining the services and path forward towards further specification of the EarthCube Architecture. This workshop is expected to take place within a few months of the TO hire.

Science Community RCN: This RCN is in the planning stages. It aims to create a network of interdisciplinary researchers, who will target the concept of cross-domain requirements and paths forwards towards full implementation. The role of governance or lack of this is currently being explored by the LC. This activity is seeking support/participation from the ESSO.

Expectations and Metrics of Success & Readiness

EarthCube and its various components (including funded projects and governance) still needs to establish a clear definition of success for its overall performance and impact. The metrics that can be used to measure and evaluate progress need to be developed. Initial activities include:

- As an indicator of EarthCube’s overall impact, the LC will request the ESSO to compile a list of peer-reviewed papers that have been published and attributed in annual reports of projects supported by EarthCube and their citations.
- Productive collaborations between geoscientists and computer scientists (and other technology experts) can also be viewed as a major success of EarthCube, but is somewhat difficult to quantify. The Engagement Team is actively looking at collaboration network models to try to assess the impact that EC has had on generating collaborations across this “gap” and illustrating the extent of these collaborations. (e.g. a graph or network analysis showing institutions and individuals that have collaborated on projects and how these have changed as the result of EC).
- The LC plans to write one or more high-profile papers that highlight some of the exciting innovations that have been prototyped on two-year funded projects and make them accessible or intelligible to geoscientists.

We also plan to evaluate separately (1) the success of EC as compared to any other similar-sized (budget-wise) NSF funding initiative, as measured by publications, citations and other products and (2) the success of EC in connecting the numerous products from individual projects via a common architecture.

The LC has emphasized to the ESSO the urgent need to showcase EarthCube achievements and progress on the EC web site. Products that have been generated by EarthCube projects often lack EarthCube branding and are not visible as a collection. We are actively encouraging new and existing projects to highlight their association with EarthCube.

The Path Forward

The LC believes significant progress can be made by beginning to centralize some EC social and technical capability at the ESSO. The TAC and CDF recommend a phased implementation of specific services that are needed to develop the first components of the EarthCube system that will allow integration of capabilities in the funded projects and the members of the Council of Data Facilities and initial use by geoscientists. *The LC proposes developing the EC system in an Agile manner, implementing key components first, identifying additional needs, and iterating toward an EC architecture that meets the needs of the EC scientific community.*

The LC recommends the transition of ESSO to a Project Office with the necessary resources to ensure the successful advancement of the EarthCube program. This includes the appointment of the Technical Officer and we seek to establish the position of an EarthCube Executive Director to provide a dedicated, high-level person to take on day-to-day decisions and communication needed to better prepare the community for the needs of fostering partnerships, integrating activities and responding to a growing number of geoscience participants.