Geophysical Research Abstracts
Vol. 15, EGU2013-6523, 2013
EGU General Assembly 2013

Assessing readiness of cyberinfrastructure resources for cross-domain interoperability: a view from an NSF EarthCube roadmap

Ilya Zaslavsky (1), Alva Couch (2), Stephen Richard (3), David Valentine (1), Kerstin Lehnert (4), Karen Stocks (1), and Philip Murphy (5)

(1) UCSD, San Diego Supercomputer Center, La Jolla, United States (zaslavsk@sdsc.edu), (2) Tufts University, Boston, United States, (3) Arizona Geological Survey, Tucson, United States, (4) Columbia University, Lamont-Doherty Earth Observatory, Palisades, United States, (5) The Redlands Institute, Redlands, United States

EarthCube is a new research initiative of the U.S. National Science Foundation, with the mission to develop community-guided cyberinfrastructure integrating data, models and other resources across geoscience disciplines. Analysis and modeling of physical processes that are not confined to disciplinary or jurisdictional boundaries, requires that data and models can be re-used outside the original context in which they are collected or developed. Infrastructure readiness for cross-domain interoperability encompasses the capabilities that need to be in place to enable such information re-use and ensure that it is both scientifically sound and technically feasible.

In an ideal cross-domain information integration scenario, resources can be discovered via federated catalogs and linked nomenclatures from neighboring domains, while standard data services can be used to transparently compile composite data products and to integrate information using commonality in key data characteristics related to shared models of spatial features, time measurements, and observations. The main premise of the cross-domain readiness assessment is that when accessing domain resources from another domain, a user expects to be able to discover these resources, interpret them, retrieve the information, and integrate it with other data. Documentation of the resource must be sufficient for a user in a different context to determine fitness for use, and establish trust in scientific soundness.

As part of an EarthCube roadmap focused on cross-domain interoperability, we explored a number of approaches to cyberinfrastructure readiness assessment, addressing both readiness of existing resources, and readiness of processes that enable cross-domain communication and information exchange across disciplinary boundaries. Our initial assessment considers basic infrastructure components required to enable cross-domain interoperability in the geosciences. These components, and the evaluation metrics, include:

1) Metadata catalogs, at the appropriate community defined granularity, that provide standard discovery services over datasets and other resources of the domain. Metrics: use of specific metadata content standard, as well as catalog search capabilities (availability of a web page interface, API, standards-compliant API such as OGC CSW) and harvest capabilities (whether harvesting is supported, and the API used).

2) Vocabularies that support unambiguous interpretation of domain resources and metadata. Metrics: use of controlled vocabularies, their documentation and standards-compliance (e.g. SKOS, OWL), and type of online access provided.

3) Services or other mechanisms used to access data repositories and other resources including models, visualizations and workflows. Metrics: types of information access, query, processing and visualization capabilities, and their standards-compliance.

4) Formal information models defining structure and semantics of the accessed information. Metrics: is there a formal model, how is it documented, its standards-compliance, and data instance validation processes.

Additional factors evaluated included assessment of community governance processes, and use of persistent identifiers.

The presentation will describe the details of the cross-domain readiness model as enabling effective com-

munication among scientists, governance bodies, and information providers, and include an initial readiness assessment and a cross-domain connectivity map for the geosciences. In addition, we demonstrate “fitness-for-use” workflows implementing cross-domain research scenarios, which are designed to collect user’s data integration experiences and make them available via searchable catalogs.