

[EarthCube - Cross-domain Interop](#) [1]

Submitted by superadmin on Fri, 2012-06-29 19:51 Thursday, July 19, 2012 - 10:30 to 12:00

Event: [Summer Meeting 2012](#) [2]

Session Type: [Breakout](#) [3]

Media/Video: [slides and google doc of notes](#) [4]

[link to earthcube session](#) [5]

[session recording \(streaming\)](#) [6]

[session recording \(download\)](#) [7]

Expertise Level: [Intermediate](#) [8]

Identifier: doi:10.7269/P38G8HMP

Collaboration Area: [Information Technology and Interoperability](#) [9]

Abstract/Agenda:

The goal of cross-domain interoperability is to enable reuse of data and models outside the original context in which these data and models are collected and used and to facilitate analysis and modeling of physical processes that are not confined to disciplinary or jurisdictional boundaries. A new research initiative of the U.S. National Science Foundation, called EarthCube, is developing a roadmap to address challenges of interoperability in the earth sciences and create a blueprint for community-guided cyberinfrastructure accessible to a broad range of geoscience researchers and students.

In this breakout session, we will report on the EarthCube roadmap development from the perspective of cross-domain interoperability. A discussion of several cross-domain use cases, challenges and solutions explored in the EarthCube interoperability project will be presented. In particular, we will focus on (1) readiness of domain information systems for cross-domain data and model re-use, (2) the initial readiness assessment, (3) determining EarthCube development priorities given a multitude of use cases, and (4) involving a wider community of geoscientists in the development and curation of cross-domain resources.

Notes:

[Earthcube.ning.com](http://earthcube.ning.com)

<http://earthcube.ning.com/events/interop-session-at-the-esip-federation-meeting> [5]

I. Zaslavsky

- Four groups – governance, workflow, semantics and ontologies, data discovery, mining, and integration web services
- The structure of Earthcube has changed – this was the initial set-up and structure
- X-Domain interop (this talk) –
- Have 66 members – not all there for development of roadmap – lots of webex and emails
 - Represented by many earth domains
- What does cross-domain mean
 - Issues is collected by many different times and places – used out of the context that they were initially intended to be used
 - Means requires information integration need systems to foster reuse of data and models
 - Not only technical solutions
- The product will never be complete because will always have incomplete metadata and always have different uses of the data
- Want several process loops – includes both bottom up and top down – how managed is controlled by governance
 - Technical options are given to decision makers
- The quality of readiness is how well this works
- Architecture
 - Not building from scratch – there are large products

- If want data from another system – need to discovery, interpret what found, access data & grab (or sub-set) and integrate with your data
- Translates into 4 components that need to be exposed
 - Catalogues, vocabularies...archives, data models
- Different clouds can be different domain – earthcube will manage cross domain resources – vocabulary crosswalks – service brokers – information model mappers
 - Need information models, following baseline information
 - If gml application schema – it will be compatible – otherwise not
 - For services – there is a set of OGC services that will help
 - Profiles for baseline
 - Still problem of how cross domain resources will be managed
- Scope (still looking at architecture)
 - Some influence on how domain follow standards and what standards will be useful
- For users to get at domain – vocab, metadata catalogs, data archives, and community data models have to be exposed or else need services
 - There are 3 tiers
 - 1) produce data
 - 2) main systems that make data available to larger community – don't community
 - 3) connect
 - ? how similar or not to NASA model – if you expose one data center to other... NASA is a box – NASA centers are a federation because they have a common data repository – but they do more – can be separate data model, but they all converge
 - NASA are top down (back in the day) – learning opportunities presented by the DAX – example of top up...
 - NASA are top down... NOAA are bottom up – it is the placement of archives along site the data (Ken)
 - Wondering if earthcube is using data archive in a different way
 - Ex. Cuahsi – have time series data – what is the responsibility of managing
 - Everything in black box is archive (Ken) – need all services
 - Dave (USGS) – these diagram parallel USGS systems – analysis for climate science center networks – these are made up of tiers
 - The data archives – may not belong – archive is actually at the bottom
- Community Inventory Page from website
 - Look at what is across domains
 - If you have some brief diagrams/compare – in dropbox (connect via website)
 - ACTION – add data to dropbox
- EarthScienc_WS_US (key domain system components – off Inventories website)
 - Also to be filled in
 - Many fields are blank
 - This is a catalogue or inventory (open for editing) can see where each system
- Readiness assessment
 - Vocabulary, catalog metadata, information model conceptual, data access API... have list of assessment
 - Went through and “scored” assessments
 - About 2000 systems
- Know what exists and what the gaps are
 - Next look at where to focus
 - Collect use cases and see what cross domain systems they uses
 - Both small and large – focus on technical challenges
- Tier 1 – science use cases – need to select use cases for developing cyberinfrastructure & provide community to implement
- Specific use case templates – provide to uses to fill in
- Hypoxia – stations from Mississippi delta and gulf – use Cuahsi
 - 7 had dissolved oxygen
 - Only 1 pointed to excel and did not have metadata or controlled vocabularies
 - No management of how data is managed as part of publication
- Developed workflow – collected searchable catalogue, collect records, look at semantic mapping

- When have data - not sure if it is compatible
- Ken (NODC) - making argument against low barrier for entry into earthcube - data in NODC archive is low quality because have low barrier - but showing iRod standards
 - Question is what low barrier really - letting anything in OR make documenting assistant
 - Low barrier is written as - through data into the cloud
 - Low barrier is a dynamic process
 - Evolving process - what works so far - is when you have good metadata
 - First define domains of ownership then what is shared across domains and then how share and then governance ways in on how it happens (Siri)
 - One of the thing that ISO has is the standard mechanism for user feedback - concept is there for structure information for use limitation (Ted) - how get feedback back to data center to improve for the next user
 - Came up with notion of "fitness for use workflow" - will provide additional annotation - what will be the schema - eventually come up with a set of annotation reusable data
 - (Ted) there are mechanism for users to record measures for quality and results of test or compliance against those datasets - need user input
- Look at models - proxies for use cases
 - Look at catalogues of models - NOAA, EPA, European
 - Look at visualization of these - see how organized, computer, domains, systems
 - No single source of models from NASA - will do this
 - Give picture of how data flows across geosciences
 - (using Pivot with some customization)
- Map of earthcube - X-domain model uses (cross domain)
 - Only one catalogue
 - Use as source of prioritization of development of earthcube
 - How confident the visualization metadata
 - From a European project - TESS - if have model - submit it
 - Some models are used one other thousands of time - need frequency of use and popularity
 - Can do this for several of the catalogues (from website)
- ACTION - if you know of model catalogues - tell them
- For gcmb - should have a list of NASA catalogue (add thoughts to google doc)
- Roadmap - common
- Have been talking about spiral development - activity that have different time spans and mechanism - how organized and governed is the question
 - Suggest interoperability institute
- Q - for institute - would it be a representative for the semantic group
 - It would be integrated
 - Q Would brokering be part of solution - yes
 - Brokering can work with any type of input
- Q (Ken) - previous session (Earthcube - discovery, mining, ...) - work with data producer - architecture here - assume archives are part of it
 - Now (new figure) - access multiple types
- Ted - 2 ways that you intrude - talk technical - but also social intrusion - this helps to facility conversion into the technology pace - benefits of uses for the system - want mind control (ish)
- Jenny(AZ geologic survey) - research on historic infrastructure - combination of both social and technical solution to interoperability problems - it is equally important
- Interoperability institute
 - List of how structure organization - real better than virtual... not all worked out yet
- Governance and GII for interoperability institute
 - Want to be managed at cross domain level
 - From science - long tail = geoscience commons
 - TED - need training and outreach → on last slide
 - Focus on geospatial standards
 - Some pieces are governed and some are not... depends on scope of earthcube
 - Ken - don't see anything as an administration layer - where is the governance that

- tracks metric, shows success, etc. → should be part of governance
- Q - cloud diagram - trying to pull in different group - whole architecture is still a research issue - want to talk about the architecture - especially the semantics
 - Vocabularies are domain specific and have semantic crosswalk
 - Dave (USGS) - diagram too small - not able to read/comment on it
- Start building the platform for earthcube & establish agreements about interfaces to create and manage the data
- System will have (in addition to components from concusses diagram)
 - Focus on documentation, curation, build cross domain interop platform and interoperability deployment stack
- Plans for summer-fall
 - MS eScience meeting - Chicago
 - AGU - several sessions
 - Proof-of concept demos
- Been a lot of discussion from the Charrette
- Can capture metadata - how capture intent and context of measurements - how capture semantic of measurement in context
 - There should be some responsibility from data publishers
 - Requires trust - using weather radar in hydrologic modeling
- Q - how combine the 2 working groups (discovery and interoperability) - governance issue - need to scope to problem
- Q - Matt Jones ECSB - governance and GII - looks like small groups contribute to large group in the center - draw diagram to show cross community
 - Workflow group (yesterday) - have similar diagram - confusion as to which group does what - each group is involved in governance
 - Lee Allison - when set up governance - have liaison from each concept team - work with committee - members are also PI of other work groups - had a number of webinars/workshop - dealt just with interactions back and forth - coming back and following Charrette - analyze governance aspects and extract - starting to follow up with groups (posted online) - this group is the leading group with governance issues
 - Asking groups what they need to operate internally and then feed back into other groups - providing services back to each group - then extract their requires to develop system to facilitates interaction between group and structure things

Actions:

if you know of model catalogues - tell them

Session Leads:

Name: [Ilya Zasklavsky](#) [10]

Organization(s): [San Diego Super Computing Center](#) [11]

Notes takers:

Name: [Kelly Monteleone](#) [12]

Organization(s): [University of New Mexico](#) [13]

Email: krbm@unm.edu [14]

Creative Common License: Creative Commons Attribution 3.0 License

Teaser: Cross-domain interop enables reuse of data/models outside the original context. Report on the #EarthCube roadmap dev & opps

Keywords: [EarthCube](#) [15]

Source URL: <http://commons.esipfed.org/node/468>

Links:

[1] <http://commons.esipfed.org/node/468>

[2] <http://commons.esipfed.org/event/summer-meeting-2012>

[3] <http://commons.esipfed.org/session-type/breakout>

- [4] <https://drive.google.com/>
- [5] <http://earthcube.ning.com/events/interop-session-at-the-esip-federation-meeting>
- [6] <https://esipfed.webex.com/esipfed/ldr.php>
- [7] <https://esipfed.webex.com/esipfed/lsr.php>
- [8] <http://commons.esipfed.org/taxonomy/term/261>
- [9] <http://commons.esipfed.org/collaboration-area/information-technology-and-interopability>
- [10] <http://commons.esipfed.org/node/467>
- [11] <http://commons.esipfed.org/taxonomy/term/317>
- [12] <http://commons.esipfed.org/node/430>
- [13] <http://commons.esipfed.org/taxonomy/term/297>
- [14] <mailto:krbm@unm.edu>
- [15] <http://commons.esipfed.org/taxonomy/term/314>