ESC as an Emergent Information (Eco)system **[1]**

Submitted by superadmin on Fri, 2012-11-30 18:39 Thursday, January 10, 2013 - 10:30 to 12:00
Event: Winter Meeting 2013 [2]
Session Type: Breakout [3]
Expertise Level: Intermediate [4]
Collaboration Area: Earth Science Collaboratory [5]
Abstract/Agenda:
Premise: The Earth Science Collaboratory, by its very nature, must have inclusiveness as a Prime Directive. The politics of top-down system architecture and design is problematic to achieving this. Instead, an ecosystem of information systems may be the way to realize the ESC. We will discuss questions like:

- What kinds of "selection pressures" favor ESC development?
- Can we design inter-species interactions to favor ESC development?
- What can we learn from Ecology, Eco-informatics and Bioinformatics about natural ecosystems that might be applicable to our info-ecosystems?
- Are there useful correspondences and metaphors, such as
    - Selection Pressure <-> Grant Requirements
    - Invasive Species <=> Disruptive Technology

Agenda:

- Chris Lynnes: Introduction
- Peter Fox: "Knowledge Networks and Science Data Ecosystems"
- Steve Young: "What Can Nature Teach Us About Improving Earth Science Data Access?"
- Discussion

Notes:

Topics

- Intro to ESC
- Peter Fox - Knowledge Networks and Science Data Ecosystems
- Steve Young - Learning From Life - What can Nature Teach Us to Improve Earth Data Access?
- Group Discussion

Presentations

- Intro
    - Intent - recap of talks about ESC ideas as ecosystem, then discuss
    - What are the parts of the ESC that can be treated as an ecosystem?
    - What are the actions we can take to enhance these ecosystems?
- Knowledge Networks (Peter Fox)
    - What is the activation function is for super nodes? What are the characteristics of activation functions? Is ESC positioned to be a "super node"?
    - (Sun-powered slide) Ways to assess ecosystem networks, and informaticists build content on the left? How do you "power" the eco-system? Make explicit the knowledge of the connections, and the models of how things fit together. A shared awareness of the relationships in the eco-system. People have a partial understanding of the interconnectedness of the network. The "person-in the middle" has complete knowledge. THere is usually not a single person, but multiple who have to negotiate a common view of the network.
    - How do you make the relationships explicit? Use-case driven approach to extract these information models.
    - Deterministic-nature of the presentation, how do you deal with stochasticity?

Middle-organizing framework is less stochastic, but the instance-data can be constantly changing, even wrong (i.e., open-world)

- Learning from Life (Steve Young)
    - The Fifth Discipline - Peter Sangee, how to build a shared vision from many people
    - Differentiate between interoperability and cooperabilty
    - Any life sciences analogies to having multiple copies or data? Ant swarm, multiple copies of an organism. In terms of efficiency, replication / redundancy for resiliency, built in redundancy. Design purpose of replication. Copies and life - population genetics, more copies of DNA you have, the faster the system evolves. Larger the population, look at in terms of code, multiple copies of algorithms more people will mutate and evolve it.
    - Shotgun sequencing, breakup the DNA, get multiple segments (smaller). Look at base pairs 1-500, 400-700, etc. Eventually samples whole sequence, Know how to take multiple sequences and piece them together, in biology, DNA alignment. Shotgun sequencing analogy, how do you piece together the whole picture from short sequences overlapping. Multiple copies of each segment, each segment has some error.
    - Shotgun sequencing is similar to how collaborative science works.
    - What is the equivalent in a network of mutation, where errors are beneficial? How is mutation introduced into science networks? Selective pressure chooses the mutations that survive. Changing projects, people leaving projects, etc. are analogous to changing mutations in science networks. Unreadability of a data product due to changing formats (or inconsistencies) is a mutation.
    - Can we design a collaboration network to promote "accidental" discoveries? Often conferences are intended to due this through social events (receptions, etc.)
    - Making the network explicit can help identify supernodes.
    - NACP database of funded project could serve as a testbed.
    - How can we develop an ESC that's as agile as species in the environment. Most species don't adapt, how do you enable failures and sacrifices for the success of the network? Diversity is important for this.
    - On the other end, what are the differences between a natural ecosystem and a technical ecosystem?
    - There's not always a one-to-one relationship between entities, in some cases, a group of entities work together to provide a "substrate" to be consumed by other entities. E.g., no one plant provides the oxygen for a species, they all provide oxygen consumed by animals.
- Group Discussion

Notes:

Chris Lynnes: Introduction

- develop a system or meta-system that allow sharing of artifacts in the earth science discipline
    - sharing of artifacts: sharing data, tools, workflows, results, and full articles
    - sharing of contextual knowledge: of each of the artifacts
    - conviction that the ESC is not something that we need to design for top down, because its very nature needs ot be highly inclusive, such as resources that didn't know they would have to come into this infrastructure
    - when you look into sociopolitical landscape, its unrealistic to think this could be top down
    - instead, we need to build it from the ground up as an emerging system
    - AGU two talks from an ECO system standpoint

1. What components do we have?
2. Which of these components have actionable items?

Peter Fox: "Knowledge Networks and Science Data Ecosystems"

- Multiple examples of Ecosystems in science
- Importance of data repositories to which eventually lead up to people, these are all complex networks
- Complex == meaning
- Semantic networks are ones where the nodes and relations are named and typed
- Interesting property: scale-free
- What is the goal of the ecosystem
    - This drives the ontology
    - Have to explore the network

Questions:

1. Go a back to “What is a knowledge network? How does it work?
    1. The problem of an emerging ecosystem is what happens in the middle

Answer:

- Making explicitly the knowledge and direction
- Making the knowledge network explicitly
- Can’t just represent single view, has to show multiple view or common view of network
- Techniques of developing information models from use cases
- What are the questions the person is asking the network

2. middle organizing framework

- have to be prepared to change that intermediate
- the high level organizing framework

Steve Young: "What Can Nature Teach Us About Improving Earth Science Data Access?"

- how can we embrace mobility
- similar pattern of data production to Moore’s Law
- the bioinformatics world had been producing the majority of data, however, the geoinformatics group is now producing more and more data
- right now it looks like we can store all of the worlds data in 4 grams of DNA
- complexity built on stunning simplicity
- can we sequence the genome of our earth data?
- Can we make our data more simple
- Can we conduct a deeper analysis of our data
- From interoperability to cooperability
- In ecosystems we see organisms competing and cooperating
- Can we analyzed our earth systems interactome
- Recommendations
    - Gather more insights from the study of life
    - Evolve our systems
- strive for cooperability and simplicity
- interact with bioinformaticians; joint workshops?
- Bioinformaticians are ahead of us right now

Lessons from life:

- be mobile; keep moving
- keep it simple

- Looking at system and data
- Multiple dimensions form looking at super organisms kinds of concepts to we don't generally have one copy, we have populations of organism, we don't like to have N number of copies of the same data
- The interesting thing about copying in this context is the error, which can be beneficial, which is the driver for evolution.
- Another way to think about it could be for an interaction to take place between two nodes who haven't interacted before would be for them, they need to accept new nodes and spontaneous interactions
- Collaborations from interactions and accidental interactions, and serendipity

ESC aspires to become somewhat of a supernode

- do we see metrics that show that we're moving in the right direction
- agile to our technology
- the diversity part is clearly part of the answer, because some of the organisms

We could also think about censor data, people with smart phones, what can we make out of all of these people generating data, how can we leverage that

Session Leads:

Name: Christopher Lynnes [6]
Organization(s): NASA Goddard Space Flight Center [7]
Email: christopher.s.lynnes@nasa.gov [8]

Presenters:

Name: Peter Fox [9]
Organization(s): Rensselaer Polytechnic Institute [10]

Name: Steve Young [11]

Notes takers:

Name: Eric Rozell [12]
Organization(s): Rensselaer Polytechnic Institute [10]

Name: Angela P. Murillo [13]
Email: amurillo@email.unc.edu [14]

Creative Common License: Creative Commons Attribution 3.0 License
Teaser: How can an ecosystem of information systems and resources realize the vision of the Earth Science Collaboratory?
Keywords: collaboratory ecosystem emergence [15]

Source URL: https://commons.esipfed.org/node/759

Links
[1] https://commons.esipfed.org/node/759
[2] https://commons.esipfed.org/taxonomy/term/464

[3] https://commons.esipfed.org/session-type/breakout
[4] https://commons.esipfed.org/taxonomy/term/261
[5] https://commons.esipfed.org/collaboration-area/earth-science-collaboratory
[6] https://commons.esipfed.org/node/542
[7] https://commons.esipfed.org/taxonomy/term/246
[8] mailto:christopher.s.lynnes@nasa.gov
[9] https://commons.esipfed.org/node/327
[10] https://commons.esipfed.org/taxonomy/term/221
[11] https://commons.esipfed.org/node/1245
[12] https://commons.esipfed.org/node/292
[13] https://commons.esipfed.org/node/633
[14] mailto:amurillo@email.unc.edu
[15] https://commons.esipfed.org/taxonomy/term/583