

Session Title: IN54A. Advancing Partnerships, Collaborative Platforms, and Knowledge Networks in the Earth Sciences II

Session Type: Oral

Session Topic: Earth and Space Science Informatics (IN)

Convener: Dawn J Wright

Convener: Sylvia Murphy

Location: 2020 (Moscone West)

Co-Sponsor(s): (none)

Index Term 1: (none)

Index Term 2: (none)

Index Term 3: (none)

Index Term 4: (none)

ID of primary accepted session entered into 'notes to Admin': (none)

Index Term 1: [1908] INFORMATICS / Cyberinfrastructure

Index Term 2: [1938] INFORMATICS / Knowledge representation and knowledge bases

Index Term 3: [1960] INFORMATICS / Portals and user interfaces

Index Term 4: [1976] INFORMATICS / Software tools and services

Session Merged Into: (none)

Co-Sponsor(s): Atmospheric and Space Electricity (AE), Atmospheric Sciences (A), Biogeosciences (B), Cryosphere (C), Earth and Planetary Surface Processes (EP), Education (ED), Geodesy (G), Geomagnetism and Paleomagnetism (GP), Global Environmental Change (GC), Hydrology (H), Mineral and Rock Physics (MR), Natural Hazards (NH), Near Surface Geophysics (NS), Nonlinear Geophysics (NG), Ocean Sciences (OS), Paleoceanography and Paleoclimatology (PP), Planetary Sciences (P), SPA-Aeronomy (SA), SPA-Magnetospheric Physics (SM), SPA-Solar and Heliospheric Physics (SH), Seismology (S), Societal Impacts and Policy Sciences (SI), Study of Earth's Deep Interior (DI), Tectonophysics (T), Volcanology, Geochemistry, and Petrology (V)

CONTROL ID: 1465384

TITLE: Partnerships Drive Informatics Solutions for Biological Imaging at Ocean Observatories
(Invited)

AUTHORS (FIRST NAME, LAST NAME): Heidi M Sosik¹, Joseph Futrelle², Andrew R Maffei³

INSTITUTIONS (ALL): 1. Biology, Woods Hole Oceanographic Institution, Woods Hole, MA, United States.

2. Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA, United States.

3. Computer and Information Services, Woods Hole Oceanographic Institution, Woods Hole, MA, United States.

ABSTRACT BODY: In the big-data, era informatics-oriented partnerships are needed to achieve improved scientific results and understanding. Our teams' experience shows that formal methodologies to build interdisciplinary partnerships enable us to efficiently produce needed technological innovation. One-on-one partnerships between individual research scientists and informaticists provide a crucial building block for supporting larger, nested partnerships. We present one such partnership as an example.

As ocean observatories mature, they produce data at a pace that threatens to overwhelm the capacity of individual researchers to manage and analyze it. Our multi-disciplinary team has addressed these challenges in the context of a study involving very large numbers (~1 billion) of images collected by Imaging FlowCytobot, an automated submersible flow cytometer that continuously images plankton at up to 10hz. These data provide novel insights into coastal ecosystem dynamics, including characterization of biological responses to environmental change and early warning of harmful algal blooms.

In contrast with the traditional focus on technology adoption, we have instead emphasized building partnerships between oceanographers and computer scientists. In these partnerships we identify use cases, design solutions, develop prototypes, and refine them until they meet oceanographers' science needs. In doing so we have found that rapid and significant advances do not always require technological innovations, but rather effective communication, focus on science outcomes, and an iterative design and evaluation process. In this work we have adopted a methodology developed in the Tetherless World Constellation at Rensselaer Polytechnic Institute, a framework that has been used for several data-intensive earth science applications.

The prototype system produced for Imaging FlowCytobot data provides simple and ubiquitous access to observational data and products via web services and includes a data dashboard (<http://ifcb-data.whoi.edu/>) that enables near-real-time browsing of images. Data can now be reprocessed with improved algorithms in a fraction of the time (weeks now, instead of years). Public web services replace researchers' need to gain access to internal networks or custom software. Links to image data can now be cited in publications and other documentation, emailed, posted to social networks, etc. A key strategy has been to enable these new capabilities without disrupting existing, working systems (e.g., no requirement to reformat existing data or rewrite analysis codes). The new data system is currently in use by multiple researchers deploying Imaging FlowCytobots and providing input that is informing continued development.

<http://ifcb-data.whoi.edu/>

KEYWORDS: [1976] INFORMATICS / Software tools and services, [1996] INFORMATICS / Web Services, [4855] OCEANOGRAPHY: BIOLOGICAL AND CHEMICAL / Phytoplankton, [4894] OCEANOGRAPHY: BIOLOGICAL AND CHEMICAL / Instruments, sensors, and techniques.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material: 0

Contact Details

CONTACT (NAME ONLY): Heidi Sosik

CONTACT (E-MAIL ONLY): hsosik@whoi.edu

TITLE OF TEAM:

CONTROL ID: 1488157

TITLE: Knowledge Networks and Science Data Ecosystems. (*Invited*)

AUTHORS (FIRST NAME, LAST NAME): Peter Arthur Fox¹

INSTITUTIONS (ALL): 1. Tetherless World Constellation, Rensselaer Polytechnic Inst., Troy, NY, United States.

ABSTRACT BODY: In an era where results from inter-disciplinary science collaborations are widely sought after for assessment reports, and often policy development and decision making, the prospect of synthesizing and interpreting complex data from myriad sources has suddenly become daunting. Even more demanding is the increased need to explain science analysis results to non-specialists, or answer their questions. These multi-stakeholder networks are often poorly understood, or documented.

Recent network developments for an NSF-funded Data Interoperability Network project (Integrated Ecosystem Assessments for Marine Ecosystems) have highlighted the importance of formally characterizing the network of people, organizations (together these are stakeholders), resources, relationships, etc. in addition to the data and information networks.

Each stakeholder in a network (in particular the marine ecosystem community, broadly defined) is a repository of knowledge about her or his domain. Too often this knowledge is ‘grey’ (tacit) and not accessible in a way that questions of interest can be formulated, posed, answered and assessed. Knowledge networks provide representations of a look into a knowledge base with the goal of gaining insight and understanding into various attributes of a real network. A key aspect is that the relationships among the things in the network (e.g. Organization A has a memorandum of understanding with Organization B for personnel exchange, or Person B is director of Organization A and an advisory board member for Organization B). Simpler examples of knowledge networks, where there is only one or a few simple (less well defined relationships), are co-authorship networks in peer reviewed publication, or friends in a social network. The knowledge networks we seek here are richer and necessarily more complex.

In this contribution, we present an approach to model such knowledge networks and discuss how they may begin to address the questions of the non-specialist in an era of Big Data.

KEYWORDS: [1918] INFORMATICS / Decision analysis, [1930] INFORMATICS / Data and information governance, [1938] INFORMATICS / Knowledge representation and knowledge bases, [1934] INFORMATICS / International collaboration.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material: 0

Contact Details

CONTACT (NAME ONLY): Peter Fox

CONTACT (E-MAIL ONLY): pfox@cs.rpi.edu

TITLE OF TEAM:

CONTROL ID: 1494101

TITLE: Mind the Gap: furthering the development of an international collaboration in marine data management (*Invited*)

AUTHORS (FIRST NAME, LAST NAME): Helen M Graves¹, Stephen P Miller², Roger Proctor³, Dick Schaap⁴

INSTITUTIONS (ALL): 1. Information Management, British Geological Survey, Nottingham, United Kingdom.

2. Scripps Institution of Oceanography, University of California, San Diego, CA, United States.

3. IMOS, University of Tasmania, Hobart, TAS, Australia.

4. MARIS, Voorburg, Netherlands.

ABSTRACT BODY: A large and ever increasing amount of marine data is available throughout Europe, USA, Australia and beyond. The challenges associated with the acquisition of this data mean that the cost of collection is high and the data itself often irreplaceable. At a time when the demand for marine data is growing while financial resources for its collection are being dramatically reduced the need to maximise its re-use is becoming a priority for marine data managers.

A number of barriers to the re-use of marine data currently exist due to the various formats, standards, vocabularies etc. used by the organisations engaged in collecting and managing this data. These challenges are already being addressed at a regional level by projects in Europe (Geo-Seas, SeaDataNet etc.), USA (R2R) and Australia (IMOS). To expand these projects further and bridge the gap between these regional initiatives the Ocean Data Interoperability Platform (ODIP) will establish a collaborative platform which will facilitate the development of a common approach to marine data management. Proactive dissemination of the outcomes and products of this project will promote adoption of the common standards and practices developed by the ODIP project to other organisations and regions beyond the 20 original consortium partners.

To demonstrate this coordinated approach several joint prototypes will be developed to test and evaluate potential solutions for solving the marine data management issues identified within the different marine disciplines. These prototypes will also be used to illustrate the effective sharing of data across scientific domains, organisations and international boundaries through the development of common practices and standards in marine data management.

KEYWORDS: [1916] INFORMATICS / Data and information discovery, [1912] INFORMATICS / Data management, preservation, rescue.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material: 10% at European Geosciences Union

General Assembly 2012
Vienna, Austria, 22 – 27 April 2012

Contact Details

CONTACT (NAME ONLY): Helen Glaves

CONTACT (E-MAIL ONLY): hmg@bgs.ac.uk

TITLE OF TEAM:

CONTROL ID: 1490237

TITLE: The Virtual Learning Commons: Supporting the Fuzzy Front End of Scientific Research with Emerging Technologies

AUTHORS (FIRST NAME, LAST NAME): Deana D Pennington¹, Aida Gandara^{2, 1}, Ivan Gris¹

INSTITUTIONS (ALL): 1. CyberShARE Center of Excellence, University of Texas at El Paso, El Paso, TX, United States.

2. Computer Science, University of Texas at El Paso, El Paso, TX, United States.

ABSTRACT BODY: The Virtual Learning Commons (VLC), funded by the National Science Foundation Office of Cyberinfrastructure CI-Team Program, is a combination of Semantic Web, mash up, and social networking tools that supports knowledge sharing and innovation across scientific disciplines in research and education communities and networks. The explosion of scientific resources (data, models, algorithms, tools, and cyberinfrastructure) challenges the ability of researchers to be aware of resources that might benefit them. Even when aware, it can be difficult to understand enough about those resources to become potential adopters or re-users. Often scientific data and emerging technologies have little documentation, especially about the context of their use. The VLC tackles this challenge by providing mechanisms for individuals and groups of researchers to organize Web resources into virtual collections, and engage each other around those collections in order to a) learn about potentially relevant resources that are available; b) design research that leverages those resources; and c) develop initial work plans. The VLC aims to support the “fuzzy front end” of innovation, where novel ideas emerge and there is the greatest potential for impact on research design. It is during the fuzzy front end that conceptual collisions across disciplines and exposure to diverse perspectives provide opportunity for creative thinking that can lead to inventive outcomes.

The VLC integrates Semantic Web functionality for structuring distributed information, mash up functionality for retrieving and displaying information, and social media for discussing/rating information. We are working to provide three views of information that support researchers in different ways:

1. Innovation Marketplace: supports users as they try to understand what research is being conducted, who is conducting it, where they are located, and who they collaborate with;
2. Conceptual Mapper: supports users as they organize their thinking about their own and related research;
3. Workflow Designer: supports users as they generate task-level analytical designs and consider data/methods/tools that could be relevant.

This presentation will discuss the innovation theories that have informed design of the VLC, hypotheses about the use of emerging technologies to support the process of innovation, and will include a brief demonstration of these capabilities.

KEYWORDS: [1974] INFORMATICS / Social networks, [1970] INFORMATICS / Semantic web and semantic integration, [1916] INFORMATICS / Data and information discovery, [1968] INFORMATICS / Scientific reasoning/inference.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material: 10%, Annual Meeting of the Association of American Geographers, April 2012.

Contact Details

CONTACT (NAME ONLY): Deana Pennington

CONTACT (E-MAIL ONLY): ddpennington@utep.edu

TITLE OF TEAM:

CONTROL ID: 1491085

TITLE: The iPlant Collaborative: A model for a collaborative science cyberinfrastructure.

AUTHORS (FIRST NAME, LAST NAME): Dan Stanzione¹

INSTITUTIONS (ALL): 1. Texas Advanced Computing Center, University of Texas at Austin, Austin, TX, United States.

ABSTRACT BODY: Cyberinfrastructure (CI) that supports collaboration and sharing of large datasets is of increasing importance to all areas of science, and particularly in data driven sciences. In the geosciences, the EarthCube initiative is focused on creating a new CI that allows for data driven collaborations.

The life sciences have already undertaken a similar effort, in the form of the iPlant Collaborative. iPlant is a comprehensive CI for plant sciences (and now expanding into animals and microbes) that supports large scale simulation, data analysis, workflows, and data sharing. In this talk, I will describe the iPlant Collaborative as a model for building CI for any data intensive scientific domain. The talk will cover the architecture and major components of the iPlant CI, and discuss how users are employing the infrastructure to get science done. In particular, the talk will cover what the lessons learned are from this project, and discuss what components could be reusable in geosciences and other domains, and what parts may need to be reinvented.

KEYWORDS: [1908] INFORMATICS / Cyberinfrastructure, [1932] INFORMATICS / High-performance computing, [1996] INFORMATICS / Web Services.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material: 50

Contact Details

CONTACT (NAME ONLY): Dan Stanzione

CONTACT (E-MAIL ONLY): dan@tacc.utexas.edu

TITLE OF TEAM:

CONTROL ID: 1478823

TITLE: The Community Surface Dynamics Modeling System: Experiences on Building a Collaborative Modeling Platform (*Invited*)

AUTHORS (FIRST NAME, LAST NAME): Irina Overeem¹, Eric Hutton¹, Albert Kettner¹, Scott Dale Peckham¹, James P Syvitski¹

INSTITUTIONS (ALL): 1. Inst Arctic & Alpine Research, Univ Colorado, Boulder, CO, United States.

ABSTRACT BODY: The Community Surface Dynamics Modeling System – CSDMS- develops a software platform with shared and coupled modules for modeling earth surface processes as a community resource. The framework allows prediction of water, sediment and nutrient transport through the landscape and seacape. The underlying paradigm is that the Earth surface we live on is a dynamic system; topography changes with seasons, with landslides and earthquakes, with erosion and deposition. The Earth Surface changes due to storms and floods, and important boundaries, like the coast, are ever-moving features. CSDMS sets out to make better predictions of these changes. Earth surface process modeling bridges the terrestrial, coastal and marine domains and requires understanding of the system over a range of time scales, which inherently needs interdisciplinarity. Members of CSDMS (~830 in July 2012) are largely from academic institutions (~75%), followed by federal agencies (~17%), and oil and gas companies (~5%). Members and governmental bodies meet once annually and rely additionally on web-based information for communication. As an organization that relies on volunteer participation, CSDMS faces challenges to scientific collaboration. Encouraging volunteerism among its members to provide and adapt metadata and model code to be sufficiently standardized for coupling is crucial to building an integrated community modeling system. We here present CSDMS strategies aimed at providing the appropriate technical tools and cyberinfrastructure to support a variety of user types, ranging from advanced to novice modelers. Application of these advances in science is key, both into the educational realm and for managers and decision-makers. We discuss some of the implemented ideas to further organizational transparency and user engagement in small-scale governance, such as advanced trackers and voting systems for model development prioritization through the CSDMS wiki. We analyzed data on community contributions and novice user engagement and evaluate the effectiveness of CSDMS' strategies toward these two challenges over the first 5 years based on member and user data, surveys, computing logs and web log analysis. Analysis shows that sponsored member participation in annual meetings (~30%) is relatively high. Direct CSDMS governance relies on ~4% of members. About 15% of members contributed code and metadata, and 18% use the common supercomputing resources. Technological development and documentation lie predominantly in hands of funded members, and a small number of others (~3% together). Potential new users are trained in clinics and courses, and on a one-to-one basis with quantified positive effects on self-efficacy and recruitment of new advanced developers.

KEYWORDS: [0545] COMPUTATIONAL GEOPHYSICS / Modeling, [1847] HYDROLOGY / Modeling, [1902] INFORMATICS / Community modeling frameworks, [1976] INFORMATICS / Software tools and services.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material: some of these statistics have been presented in 2012 paper to Environmental Modeling and Software. New data mining and statistics on community efforts for 2012.

Contact Details

CONTACT (NAME ONLY): Irina Overeem

CONTACT (E-MAIL ONLY): irina.overeem@colorado.edu

TITLE OF TEAM:

CONTROL ID: 1500210

TITLE: Climate-Change Problem Solving: Structured Approaches Based on Real-World Experiences

AUTHORS (FIRST NAME, LAST NAME): Richard B Rood¹, Laura J. Briley¹, Daniel A. Brown¹

INSTITUTIONS (ALL): 1. AOSS 2455 Hayward St, University of Michigan, Ann Arbor, MI, United States.

ABSTRACT BODY: Nearly two decades of experience using both seasonal and long-term climate model projections has led to the identification of a set of characteristics of the successful use of climate knowledge in planning and adaptation applications. These characteristics include end-to-end knowledge systems, co-generation or co-production of solution approaches by scientists and practitioners, and tailoring climate model information to the decision-making processes of the specific application.

Glisacclimate.org strives to apply the growing body of research into the successful use of climate knowledge using a set of prototype, real-world applications. We describe an online problem-solving environment whose design is based on the characteristics of the successful use of climate predictions and projections by practitioners such as resource managers, urban planners, public health professionals, and policy makers. Design features of Glisacclimate.org include:

Based on principles extracted from social science studies of the use of climate information.

Anchored on structured templates of problem solving with the identification of common steps in problem solving that are repeated in one application to the next.

Informed by interviews with real-world users who desire to incorporate climate-science knowledge into their decision making.

Built with open-source tools to allow participation of a community of developers and to facilitate the sustainability of the effort.

A structured approach to problem solving is described by four functions of information management. At the foundation of problem solving is the collection of existing information, an inventory stage. Following the collection of the information there are analysis and evaluation stages. In the analysis stage interfaces are described and knowledge gaps are identified. The evaluation stage assesses the quality of the information and the relevance of the information to the specific attributes of the problem. The development of plans and solution strategies follows from the synthesis of information.

The goals of Glisacclimate.org include the accelerated use of science-based climate knowledge in decision making. Especially in the building of the information inventory, we assert that there is substantial reuse and sharing of resources from one problem to the next. This knowledge base is captured by Glisacclimate.org as the community participating in problem solving tags information with search terms and usability descriptors through the combination of a defined vocabulary and user-defined tagging. Community members can tag particular resources found to be especially valuable; for example, the identification of a review document that is broadly relevant to Great Lakes cities. Glisacclimate.org supports evaluation and synthesis of information by linking narrative descriptions to

particular resources so that future users of a resource can immediately benefit from the previous experience of others. Both public and private spaces are provided for the development of documents.

Glisaclimate.org <http://glisaclimate.org/>

Tutorials: <http://glisaclimate.org/video-tutorials>

Who is glisaclimate.org built for? <http://glisaclimate.org/node/114>

KEYWORDS: [1938] INFORMATICS / Knowledge representation and knowledge bases, [1908] INFORMATICS / Cyberinfrastructure.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material: 0

Contact Details

CONTACT (NAME ONLY): Richard Rood

CONTACT (E-MAIL ONLY): rbrood@umich.edu

TITLE OF TEAM:

CONTROL ID: 1502662

TITLE: OGC Collaborative Platform undercover

AUTHORS (FIRST NAME, LAST NAME): Greg Buehler¹, David K Arctur², Luis E Bermudez³

INSTITUTIONS (ALL): 1. Open Geospatial Consortium, Bloomington, IN, United States.
2. University of Texas at Austin, Austin, TX, United States.
3. Open Geospatial Consortium, Rockville, MD, United States.

ABSTRACT BODY: The mission of the Open Geospatial Consortium (OGC) is to serve as a global forum for the collaboration of developers and users of spatial data products and services, and to advance the development of international standards for geospatial interoperability. The OGC coordinates with over 400 institutions in the development of geospatial standards. OGC has a dedicated staff supported by a Collaborative Web Platform to enable sophisticated and successful coordination among its members.

Since its origins in the early 1990s, the OGC Collaborative Web Platform has evolved organically to be the collaboration hub for standards development in the exchange of geospatial and related types of information, among a global network of thousands of technical, scientific and management professionals spanning numerous disparate application domains. This presentation describes the structure of this collaboration hub, the relationships enabled (both among and beyond OGC members), and how this network fits in a broader ecosystem of technology development and information standards organizations.

KEYWORDS: [1960] INFORMATICS / Portals and user interfaces, [1936] INFORMATICS / Interoperability, [1996] INFORMATICS / Web Services, [1982] INFORMATICS / Standards.

(No Image Selected)

(No Table Selected)

Additional Details

Previously Presented Material:

Contact Details

CONTACT (NAME ONLY): Greg Buehler

CONTACT (E-MAIL ONLY): gbuehler@opengeospatial.org

TITLE OF TEAM:
