

Federation of Earth Science Information Partners  
**Session & Poster Descriptions for Winter Conference 2015**



Earth Science and Data in Support of Food Resilience:  
Climate, Energy and Water Nexus

January 6-8, 2015  
Renaissance DuPont Circle Hotel  
Washington, DC

## Table of Contents

Tuesday, January 6 - Session Descriptions .....	3
<i>Breakout Sessions, 1:30 - 3:00</i> .....	3
<i>Breakout Sessions, 3:30 - 5:00</i> .....	5
Wednesday, January 7 - Session Descriptions .....	6
<i>Breakout Sessions, 1:30-3:00</i> .....	6
<i>Breakout Sessions, 3:30-5:00</i> .....	9
Thursday, January 8 - Session Descriptions.....	10
<i>Breakout Sessions, 8:30-10:00</i> .....	10
<i>Breakout Sessions, 10:30-12:00</i> .....	12
<i>Breakout Sessions, 1:30-3:00</i> .....	14
<i>Breakout Sessions, 3:30-5:00</i> .....	16
Poster and Demo Abstracts .....	18

## **Tuesday, January 6 - Session Descriptions**

### **Breakout Sessions, 1:30 - 3:00**

#### **Progress in Data Management Planning**

Room: Dupont

Session Lead(s): Nancy Ritchey

Description: Recent focus on data initiatives and policies from the government, funding agencies and publishers has highlighted the importance of Data Management Planning. How have we used this increased attention to improve our Data Management Planning policies, approaches, and tools? The goal of this session is sharing leading practices, approaches and tools that will further improve Data Management Planning across the Earth Science Data Partners. Speakers from USGS, USDA, NASA and NOAA will present their organization's perspective on this topic followed by a panel discussion addressing challenges, gaps and potential solutions.

Session Type: Breakout

Topic tags: data management plans (DMP)

Link: <http://commons.esipfed.org/node/2697>

#### **Improving Performance for Data Access Web Services**

Room: Foggy Bottom

Session Lead(s): Dave Fulker, James Gallagher

Description: This panel discussion will feature experts on data retrieval via Web services, emphasizing technologies and methodologies that will help OPeNDAP and OGC web service users (e.g., NASA/ESDIS) achieve sub-second response times for important aspects of its data discovery and retrieval services. The panel will present material, answer questions from the audience and note suggestions from the audience. The format of the discussion will be informal and invite participation from all attendees. Likely technologies to be discussed are data organization techniques, data replication, data storage technologies that are not based on file systems (array databases; cloud storage technologies) and user interfaces for on-the-fly data aggregation.

Session Type: Panel

Topic tags: Web Services, high performance, optimization

Link: <http://commons.esipfed.org/node/2714>

#### **Bridging the resolution gap between satellite data and agricultural applications**

Room: Mt. Vernon

Session Lead(s): Bill Teng

Description: These two sessions, sponsored by the Agriculture and Climate Cluster, will address the spatial resolution gap between what satellite data can offer and what agricultural applications need, which is often dictated by station observational data over small geographic areas. Focus will be on the USDA Regional Climate Hubs and the needs of their user communities, on other agricultural applications of satellite data, and on a system for the critical co-location of satellite and ground measurements. A panel/group discussion and Q&A towards the end of the second session will provide ample opportunity to bring together the presented topics, to elicit potential solutions, and for everyone to actively participate.

Session 1:

Randy Johnson, USDA FS - Climate Hubs, a national perspective

Liping Di, GMU - Agricultural applications and spatial (and temporal) resolution

Steve Kelling, Cornell U - Satellite data, bird conservation, rice field management, and resolution

Arif Albayrak, NASA GES DISC - System for co-location of satellite and ground data

Session Type: Breakout

Topic tags: USDA Climate Hubs, agriculture applications, Satellite Data, spatial resolution, SMAP, GPM

Link: <http://commons.esipfed.org/node/2705>

## **Metadata for Discoverability, Accessibility, Useability, and Understanding**

Room: New Hampshire

Session Lead(s): Ted Habermann

Description: The connection between metadata and data discovery has been strong since the early days of the Web. At the same time, metadata standards have included concepts that go significantly beyond discovery into the realms of accessibility, useability, and understanding. These elements generally do not make it onto lists of minimum metadata requirements and typically are only sparsely populated in many metadata collections. This means data that are discovered may not be accessible, useable, or understandable. As part of the Big Earth Data Initiative NASA is developing guidelines for documenting access, use, and understanding. This work is shared in the Documentation Connections section of the ESIP Wiki and will be presented during this session. The next step is developing tools and processes that will facilitate effective improvement of metadata for accessibility, use, and understanding.

Session Type: Breakout

Topic tags: documentation

Link: <http://commons.esipfed.org/node/2727>

## **Docker: open container system for developers and sys admins**

Room: Potomac

Session Lead(s): Douglas Fils

Description: Docker (<http://www.docker.com>) is popular new technology for building distributed applications. It provides a portable, lightweight way of packaging applications so they can be quickly and reliably deployed across different environments. An overview of the Docker container software and its use in various devops style patterns will be presented. The creation of containers and their deployment to remote servers will also be shown. Docker will be compared and contrasted to more common virtual machine approaches in order to better illustrate the conditions in which each are more appropriate.

Session Type: Breakout

Topic tags: docker, containers, virtualization, devops

Link: <http://commons.esipfed.org/node/2729>

## **EarthCube: A Community-Driven Organization for Geoscience Cyberinfrastructure**

Room: Private Dining Room

Session Lead(s): Anna Katz

Description: The National Science Foundation's (NSF) EarthCube initiative is a community-driven approach to building cyberinfrastructure for managing, sharing, and exploring geoscience data and information to better address today's grand-challenge science questions.

The EarthCube Test Enterprise Governance project is a two-year effort seeking to engage diverse geo- and cyber-science communities in applying a responsive approach to the development of a governing system for EarthCube. During Year 1, an Assembly of seven stakeholder groups representing the broad EarthCube community developed a draft Governance Framework. Finalized at the June 2014 EarthCube All Hands Meeting, this framework will be tested during the demonstration phase in Year 2, beginning Oct. 2014.

A brief overview of the framework: Community-elected members of the EarthCube Leadership Council will be responsible for managing strategic direction and identifying the scope of EarthCube. Three Standing Committees will also be established to oversee the development of technology and architecture, to coordinate among new and existing data facilities, and to represent the academic geosciences community in driving development of EarthCube cyberinfrastructure. An Engagement Team and a Liaison Team will support communication and partnerships with internal and external stakeholders, and a central Office will serve a logistical support function to the governance as a whole. Finally, ad hoc Working Groups and Special Interest Groups will take on other issues related to EarthCube's goals.

The Year 2 demonstration phase will test the effectiveness of the proposed framework and allow for elements to be changed to better meet community needs. It will begin by populating committees and teams, and finalizing leadership and decision-making processes to move forward on community-selected priorities including identifying science drivers, coordinating emerging technical elements, and coming to convergence on system architecture. A January mid-year review will assemble these groups to analyze the effectiveness of the framework and make adjustments as necessary. If successful, this framework will move EarthCube forward as a collaborative platform and potentially act as a model for future NSF investments in geoscience cyberinfrastructure.

Session Type: Breakout

Topic tags: EarthCube, cyberinfrastructure, NSF, Geosciences, governance

Link: <http://commons.esipfed.org/node/2704>

## **Breakout Sessions, 3:30 - 5:00**

### **What does it mean to Publish Data?**

Room: Dupont

Session Lead(s): Nancy Ritchey, Ruth Duerr

Description: Recent national and international policies on open data and research results and the increase of similar policies from research funders and publishers have motivated the Earth Science community to enable data publishing through unique identifiers. In response, data repositories are minting unique identifiers such as a Digital Object Identifier (DOI) as part of their preservation services to enable data citations. However, publishers are keenly focused on uniquely identifying the data used to create the tables and figures within the publication, not necessarily the entire data record in its original form. This begs the question, "What does it mean to publish data?". This session will bring together speakers from Academia, Funders, Data Repositories and Publishing to understand their perspectives followed by a panel discussion addressing the ambiguous definition of "data publishing".

Session Type: Breakout

Topic tags: publishing data

Link: <http://commons.esipfed.org/node/2719>

### **GEOSS Architecture Implementation Pilot Phase 7: Earth Observation Apps for end-users**

Room: Foggy Bottom

Session Lead(s): George Percivall

Description: The GEOSS Architecture Implementation Pilot (AIP) is agile development process for the GEOSS Information System. Phase-7's focus was on App development, focussing on real-world (end-user) scenario's, that use a variety of earth observation data (exposed through interoperable web services) to help the end-user make well informed decisions. In this session you will learn about the process of AIP (how to get involved), overview of the results. There will be an in-depth review of 1) the Water SBA results and 2) citizens observatories. AIP also investigated the interaction of Community Portals with the GEO Portal - this topic will also be discussed in this session. At the end the speaker will inform the audience on AIP-8 goals and how you can contribute.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/7703>

### **Bridging the resolution gap between satellite data and agricultural applications**

Room: Mt. Vernon

Session Lead(s): Bill Teng

Description: These two sessions, sponsored by the Agriculture and Climate Cluster, will address the spatial resolution gap between what satellite data can offer and what agricultural applications need, which is often dictated by station observational data over small geographic areas. Focus will be on the USDA Regional Climate Hubs and the needs of their user communities, on other agricultural applications of satellite data, and on a system for the critical co-location of satellite and ground measurements. A panel/group discussion and Q&A towards the end of the second session will provide ample opportunity to bring together the presented topics, to elicit potential solutions, and for everyone to actively participate.

Session 2:

Wade Crow, USDA ARS - Agricultural applications of SMAP and resolution

Dalia Kirschbaum, GSFC HSL - GPM applications and perspective on resolution

Panel discussion

Session Type: Breakout

Topic tags: USDA Climate Hubs, agriculture applications, Satellite Data, spatial resolution, SMAP, GPM

Link: <http://commons.esipfed.org/node/2705>

### **Metadata evaluation, consistency, compliance and improvement**

Room: New Hampshire

Session Lead(s): Ed Armstrong

Description: This session will focus on tools and approaches for the evaluation and improvement of metadata from the perspective of error, consistency and quality. Both concrete examples (existing tools and services) and abstract ideas for new services needed are encouraged.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2684>

### **Schema.org for Earth Science**

Room: Potomac

Session Lead(s): Douglas Fils

Description: This session will focus on schema.org and its applicability to Earth Science data management.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2718>

### **Birds of a Feather - Free space**

Room: Private Dining Room

Description: Birds of a feather is free, unscheduled space available if a topic emerges that you'd like to continue a conversation or work space.

Session Type: Other

Link: <http://commons.esipfed.org/node/7763>

## ***Wednesday, January 7 - Session Descriptions***

### **Breakout Sessions, 1:30-3:00**

#### **Science Software Cluster**

Room: Dupont

Session Lead(s): Nic Weber

Description: Meeting / Breakout Session for Science Software cluster.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2721>

#### **Cloud Technologies and Architectures Seminars**

Room: Foggy Bottom

Session Lead(s): Phil Yang, Thomas Huang

Description: Cloud Computing opens up new possibilities for improving Earth science. Cloud Computing also changes the way how we architect our software application in order to truly leverage the elasticity that Cloud Computing offers. This session invites speakers to present studies, technologies and architectures that fully leverage the elastic Cloud to improve our Big Earth Science data challenges.

Session Type: Breakout

Topic tags: Cloud Computing

Link: <http://commons.esipfed.org/node/7785>

### **HDF Product Designer**

Room: Foggy Bottom

Session Lead(s): Hyo-Kyung Lee, Aleksandar Jelenak

Description: Interoperable data have been a long-time goal in many scientific communities. The recent growth in analysis, visualization, and mash-up applications that expect data stored in a standardized manner has brought the interoperability issue to the fore. On the other hand, producing interoperable data is often regarded as a sideline task in a typical research team for which resources are not readily available. The HDF Group is developing a software tool aimed at lessening the burden of creating data in standards-compliant, interoperable HDF5 files. The tool, named HDF Product Designer, lowers the threshold needed to design such files by providing a user interface that combines the rich HDF5 feature set with applicable metadata conventions. Users can quickly devise new HDF5 files while at the same time seamlessly incorporating the latest best practices and conventions from their community. The HDF Product Designer can generate interoperable data in HDF5 files from the onset of their production. The tool also incorporates collaborative features, allowing team approach in the file design, as well as easy transfer of best practices as they are being developed. The current state of the tool and the plans for future development will be presented. Constructive input from session participants is always welcome.

Session Type: Breakout

Topic tags: HDF, NetCDF, CF, Cloud Computing, RESTful Services, user interface

Link: <http://commons.esipfed.org/node/7743>

### **Connecting geodata in and among governmental agencies - Compare plans submitted in response to OSTP requests**

Room: Mt. Vernon

Session Lead(s): Marshall Ma/Peter Fox

Description: The Office of Science and Technology Policy (OSTP) has taken initiatives to promote Open Data, Open Science and Open Government ( <http://www.whitehouse.gov/administration/eop/ostp/initiatives#Openness> and [http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf)). At the conclusion of the GeoData2014 workshop (<http://tw.rpi.edu/web/Workshop/Community/GeoData2014>) held in June 2014, workshop participants suggested that discussion on opening and connecting governmental data should be continued at the ESIP 2015 winter meeting. A key point discussed was that a governmental agency should avoid being a 'silo' in the development of implementation plans in response to mandates issued by OSTP. Instead, agencies should be allowed to develop individual plans then organize comment periods where all plans are vetted together and identification of common methods can be identified and exploited or ideas can be shared on implementation in individual agencies.

Some follow-on topics for discussion proposed at GeoData2014 were:

- Clear guidance (from who to who?) on how to handle exceptions to the general mandate that come out at the same time as the mandate.
- Explore mandates that are less prescriptive on implementation details are how likely they are to be more successful and longer lasting. What does a clear view of what success looks like, when does it need to be complete, and why it is valuable?
- Do mandates tied to funding tend to be more successful?
- How to link mandates to usability tests to gauge success of implementation?

- How do we identify the costs associated with NOT implementing a mandate?

We propose a pair of sessions at ESIP 2015 winter meeting: (1) Compare plans submitted in response to the OSTP requests, i.e. share content and ways to increase collaboration/coordination of the existing plans and their overall effectiveness; (2) Invite program managers, community and agency personnel for discussion on these topics.

Session Type: Panel

Link: <http://commons.esipfed.org/node/7300>

### **Global Change Information System (GCIS)**

Room: New Hampshire

Session Lead(s): Robert Wolfe

Description: The U.S. Global Change Research Program (<http://globalchange.gov>) is sponsoring the creation of a new information system, the Global Change Information System (GCIS) that provides a web based source of will provide a web based source of authoritative, accessible, usable, and timely information about climate and global change for use by scientists, decision makers, and the public. It captures and presents supporting information from the Third National Climate Assessment.

A preliminary public version of the GCIS API is available now: <http://data.globalchange.gov>

This session will present an overview of the system, status and progress with some initial information modeling and web site concepts. There will be time for discussion and feedback about the long term vision for the system.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2671>

### **Earth Science Data Analytics 101**

Room: Potomac

Session Lead(s): Steve Kempler

Description: The broad set of techniques called Earth Science Data Analytics (ESDA) has a clear meaning to everyone, though the meanings often differs depending on the various uses of the data. Data Analytics discussions can range from developing custom code for discovering the signatures in data to leveraging tools that enable predictions to be derived from heterogeneous datasets. This session, uniquely presented by field experts, attempts to introduce the scope, complexities, and possibilities presented by ESDA to further facilitate Earth science.

Guest speakers during this session will describe data analytics use cases that they employ in their work. The goal of this session is to help organize and stimulate Federation partners in thinking about how we can facilitate Earth Science Data Analytics through Information technologies and tools.

Session Type: Breakout

Topic tags: analytics data-scientist big-data techniques/methodologies

Link: <http://commons.esipfed.org/node/2724>

### **Resilience, Sustainability and Data --driven Adaptation**

Room: Private Dining Room

Session Lead(s): Gary Foley

Description: Resilience and sustainability are closely linked. Sustainability efforts tend to need to strengthen the natural capital, social capital, and economic capital so that residents and local businesses can survive, adapt, and flourish in the face of emerging trends and challenges. This session will explore how communities can align their efforts to achieve both resilience and sustainability and how earth observation data and modeling experts can inform the communities..

Speakers will be chosen who can share their perspectives on resilience and sustainability based on actual community experiences, and will engage the audience in an interactive discussion. The outcome will be insights and recommendations about policies, practices, data needs and earth observation and research opportunities that can help to create more resilient and sustainable communities. focus on the long view, developing the capacity to support the needs of a growing population in the face of climate change, resource scarcity, and other environmental

and economic pressures. Resilience efforts tend to focus on sudden events, developing the capacity to absorb and recover from shocks such as natural and anthropogenic disasters or interruptions in critical services. However, Resilience also includes focusing on trends over time that will result in the loss of needed resources or services. All of these are key considerations in the planning and management of community needs, including transportation, infrastructure, land use, energy and water supply, waste management, local commerce, job creation, poverty alleviation, and emergency preparedness. Both are also key in identifying and readying the observational data sources that can be used in the rapid analysis of sudden events and the deployment of rapid response and adaptation tools selection.

In a complex and turbulent world, unforeseen conditions can lead to unintended and cascading consequences, often undesirable ones. Therefore, resilience is a prerequisite for realization of sustainability goals and the right observational data and appropriate analysis/decision tools are a prerequisite for effective resilience. Community leaders need to strengthen the natural capital, social capital, and economic capital so that residents and local businesses can survive, adapt, and flourish in the face of emerging trends and challenges. This session will explore how communities can align their efforts to achieve both resilience and sustainability and how earth observation data and modeling experts can inform the communities..

The outcome will be insights and recommendations about policies, practices, data needs and earth observation and research opportunities that can help to create more resilient and sustainable communities.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/7756>

## **Breakout Sessions, 3:30-5:00**

### **Climate Informatics: Some use cases and future directions for an emerging domain**

Room: Dupont

Session Lead(s): Nic Weber

Description: X-Informatics is a discipline agnostic framework that offers a way to combine theoretical principles from the information, data, and computer sciences with techniques traditionally found in software engineering and systems development.

Climate informatics - which sits at the intersection of earth science disciplines (e.g. atmospheric, geologic, oceanographic, etc), computer science, and information science- is emerging as one of the most salient examples of using the x-informatics approach to address a grand science challenge. This session will offer a number of use cases that are representative of a climate informatics approach, and more generally engage in a discussion about the contribution that ESIP might make to the emerging climate informatics domain.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2720>

### **EarthCube Architecture**

Room: Foggy Bottom

Session Lead(s): Steve Richard

Description: NSF has funded three projects to develop conceptual designs for EarthCube (see <http://workspace.earthcube.org/conceptual-designs>). This session will bring together team members from the three projects to present updates on the progress of their design work, followed by an open discussion with the community to explore similarities and differences in the approaches. The purpose of the session is to identify where the approaches are convergent, to understand the rationale for differences, and to provide a forum for community feedback on design development.

This will provide important input for the projects to refine their work, as well as an opportunity for community engagement in EarthCube conceptual design.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/7758>

## **Connecting geodata in and among governmental agencies - A discussion among program managers, community and agency personnel**

Room: Mt. Vernon

Session Lead(s): Marshall Ma

Description: See 1<sup>st</sup> session description in the 1:30-3:00 section. This second session will be a discussion with program managers, community and agency personnel on these topics.

Session Type: Panel

Link: <http://commons.esipfed.org/node/7299>

## **Earth Science Data Analytics 201**

Room: Potomac

Session Lead(s): Steve Kempler

Description: The continuum of ever-evolving data management systems presents continuous challenges to the enhancement of knowledge and facilitation of science. To overcome these challenges, it is essential to understand and develop methods that enable data relationships to be examined and information to be manipulated. Come join the Earth Science Data Analytics (ESDA) cluster in our quest to decipher the different types of data analytics, generate definitions (in terms of Earth science), collect use cases, and eventually analyze tools unique to the different types of data analytics. The ultimate goal is to perform gap analysis and provide recommendations to the community. Are you interested in helping to guide the future of information analysis?

Session Type: Breakout

Topic tags: analytics data-scientist big-data techniques/methodologies

Link: <http://commons.esipfed.org/node/2723>

## **Data Needs for Energy Applications: Gaps, Traceability, Requirements**

Room: Private Dining Room

Session Lead(s): Richard Eckman, Ana Privette

Description: The ESIP Energy & Climate Workgroup aims to support the US GEO 2nd Earth Observations Assessment by engaging with the Energy private sector to identify current gaps in data, knowledge and tools. In alignment with this larger goal, the ESIP winter session aims to bring together the Energy data user community and the ESIP technical community to:

- Identify the data needs for energy related applications;
- Identify the data requirements (formats, uncertainty, accessibility and others) to ensure effective and widespread use of the data generated by the federal agencies;
- Identify best data management practices and standards to ensure that existing and future energy related datasets provide the required transparency, traceability and provenance to support private sector related applications.

Session Type: Breakout

Topic tags: Energy, Climate, Data access, use case; preservation; best practices; provenance

Link: <http://commons.esipfed.org/node/7755>

## **Thursday, January 8 - Session Descriptions**

### **Breakout Sessions, 8:30-10:00**

#### **Semantic Tech Expo**

Room: Dupont

Session Lead(s): Tom Narock, Beth Huffer

10

Description: This session will provide a forum for individuals and organizations to present their semantic technology projects and showcase capabilities through hands-on demonstrations. Following brief introductions, presenters will set up at stations where they will conduct demos and meeting attendees can test-drive applications and interact directly with the project developer to get more information.

Session Type: Workshop

Link: <http://commons.esipfed.org/node/2708>

### **ESIP in the Global Informatics Community: 3 Diverse Perspectives and Opportunities**

Room: Foggy Bottom

Session Lead(s): Genevieve Pearthree

Description: This panel-style session will present three synergistic but diverse global informatics initiatives, each of which brings a unique perspective to the conversation: Belmont Forum E-Infrastructures & Data Management Collaborative Research Action (funding agencies), COOPEUS Project (bottom-up, project-to-project), and Global Policy and Institutions: United Nations and More (international policy frameworks). This session will begin with brief presentation on each initiative, followed by panelist and then audience Q&A. This session presents an excellent opportunity for the ESIP community to learn more about, and get involved with, leading global informatics initiatives.

Belmont Forum E-Infrastructures & Data Management Collaborative Research Action: The Belmont Forum, a global coalition of science funding agencies, is supporting a 21-month effort to prioritize international funding opportunities and long-term policy recommendations to promote a more coordinated, holistic, and sustainable approach to funding and supporting global change research. This presentation will focus on preliminary recommendations, new directions based on Belmont Forum feedback, relationships to other Belmont Forum initiatives, and possible funding and policy mechanisms that might be implemented as a result of this project.

COOPEUS Project: The COOPEUS project (COOPeration EU-US) employs a bottom-up approach to establishing information and infrastructure interoperability among and between international research infrastructures across 5 environmental science domains in Europe and the United States. This presentation will outline the COOPEUS approach to overcoming the challenges of interoperability and describe some of our successes and opportunities for community engagement in these efforts.

Global Policy and Institutions: United Nations Sustainable Development Goals and More Global sustainability and development agendas will be set during 2015 and will need assessment, observations and data support. This presentation will provide a description of anticipated action and suggest ways that science data can plug into institutional processes among UN entities and other global institutions.

Session Type: Panel

Topic tags: FUNding, Belmont Forum, data management, global change research

Link: <http://commons.esipfed.org/node/2702>

### **Disaster Lifecycle I**

Room: Mt. Vernon

Session Lead(s): Emily Law, Karen Moe

Description: ESIP Disaster Lifecycle Cluster is to facilitate connections and coordinate efforts among data providers, managers and developers of disaster response systems and tools, and end-user communities within ESIP. We are establishing a testbed called Collaborative Common Operating Picture (C-COP) to facilitate sharing and validation of data products and tools that will benefit every phase of disaster lifecycle. This session will give an overview of the testbed, demonstrate use of example products and tools shared on the platform, and describe how products and tools can be contributed to C-COP. We will also hear more about the GEOSS Architecture Implementation Pilot and activities involving the GEOSS Disasters Societal Benefit area, and we can discuss opportunities for future collaborations.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2698>

## **NASA EOSDIS Evolving Technologies Discussion**

Room: New Hampshire

Session Lead(s): Katie Baynes

Description: Earth Observing System Data and Information System (EOSDIS) continues its work on a number of different projects, systems, and initiatives. This session will provide updates and facilitate discussion on a number of these activities:

- Earthdata Search and Metadata Quality (Dan Pilone)
- Common Metadata Repository Sub-second Search (Jason Gilman)
- Unified Metadata Model Status and Discussion (Katie Baynes)
- Earthdata Standards Office and the Lifecycle Process (Yonsook Enloe)
- Earthdata Code Collaborative Status and Discussion (Brett McLaughlin)
- Next Generation Application Platform Introduction and Overview (Justin Molineaux)

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2711>

## **Geoinformatics User Training: Direct Access, Live Access, Subsetting, On-line Analysis, Formats, and Conversions**

Room: Potomac

Session Lead(s): Kenton Ross, Michael Bender

Description: This session is focused on addressing issues of data access and management for novice to intermediate users who may have unique data applications with abbreviated project timelines. The session will include guided discussions on direct access (downloads), live access (e.g. opendap, thredds), subsetting, on-line analysis (e.g. giovanni), formats, and conversions, with a focus on training required to bring users up-to-speed as quickly as possible. It is the intent of this session to bring together representatives from NASA, NOAA, and USGS Data Centers, NASA's Applied Sciences' Capacity Building elements, representatives for the target user groups, and other interested parties to network, discuss, and propose possible solutions to this underrepresented issue.

Session Type: Breakout

Topic tags: Geoinformatics training, Direct Access, Live Access, subsetting, On-line Analysis, Formats, and Conversions

Link: <http://commons.esipfed.org/node/2715>

## **Preservation and Stewardship Committee reporting session**

Room: Private Dining Room

Session Lead(s): Denise J. Hills

Description: Reporting session for the Data Preservation Committee - to update everyone on our activities for the past 6 months.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2676>

## **Breakout Sessions, 10:30-12:00**

### **Drupal Working Group: Open House**

Room: Dupont

Session Lead(s): Adam Shepherd, David Bassendine

Description: Drupal stands apart from other content management solutions because of its success as an open source technology platform AND a community. In this Open House come and meet part of that community, hear about how Drupal is being used to tackle earth science objectives, and participate first hand as we seek to tackle \*your\* Drupal issues in our follow-up Code Sprint session.

Session Type: Breakout  
Topic tags: Drupal  
Link: <http://commons.esipfed.org/node/2706>

### **Birds of a Feather - Free space**

Room: Foggy Bottom  
Session Lead(s):

Description: Birds of a feather is free, unscheduled space available if a topic emerges that you'd like to continue a conversation or work space

Session Type: Other  
Link: <http://commons.esipfed.org/node/7763>

### **Disaster Lifecycle II**

Room: Mt. Vernon  
Session Lead(s): Emily Law, Karen Moe

Description: ESIP Disaster Lifecycle Cluster is to facilitate connections and coordinate efforts among data providers, managers and developers of disaster response systems and tools, and end-user communities within ESIP. We are working toward a common information architecture/model to facilitate consistent management of data products useful for disaster life cycle, and raising awareness of emerging technology that can play a disruptive role (both positive and negative) in the deployment of products and services. In addition to presentations on new technology and new data sources for disasters, we will discuss the Disasters Lifecycle cluster strategic plan for 2015 and assess activities for the coming year that will support our goals in both the testbed and the common information architecture. There will be time for discussion and feedback.

Session Type: Breakout  
Link: <http://commons.esipfed.org/node/2699>

### **New ISO 19115-1 Capabilities**

Room: New Hampshire  
Session Lead(s): Ted Habermann

Description: Two recent developments in ISO Standards bring important new capabilities to the Earth Science Community:

- (1) The ISO Metadata Standard published during 2003 (ISO 19115) was replaced during 2014 by ISO 19115-1
- (2) The data quality section of ISO 19115 was replaced by ISO 19157

The conceptual models for ISO 19115-1 and ISO 19157 are published International Standards and the XML implementations are in the final stages of development. Both will be released during 2015. This session will focus on how these new standards can help the Earth science community by addressing questions like these and others that emerge from the discussion: Can I migrate my existing metadata to 19115-1? I need to unambiguously identify metadata records in multiple repositories I need to track when changes in my metadata happen I have many existing documentation resources that can help users I have many existing web resources that can help users My datasets include measured parameters, reference and quality information My group uses local parameter names but we need standard names to share There are papers and web pages that describe the quality of my data My data quality information exists in databases or web services. Users increase our understanding of data quality. We need to keep them in the loop

Session Type: Breakout  
Topic tags: ISO Metadata Data Quality User Usecase  
Link: <http://commons.esipfed.org/node/7859>

### **Discovery Session Best Practices**

Room: Potomac

Session Lead(s): Doug Newman, Chris Lynnes

Description: The discovery cluster has dedicated most of its resources over the last few years to promoting OpenSearch as a means of discovery of earth science data. This protocol has gained a great deal of traction in the Earth Science community. So much so that a number of extensions and best practices have arose from organizations such as:

- ESIP
- CWIC
- CEOS
- OGC

We have made a concerted effort to prevent divergence of the above efforts in an attempt to promote simple aggregation of results from disparate data providers using the open search standard. This session will, primarily, focus on the current state of these extensions and best practices and what we plan to do in the future to ensure that client developers and users can exploit many different open search providers with as little custom coding as possible. We will also be discussing best practices for discovery in areas outside of open search.

Session Type: Breakout

Topic tags: discovery open search NASA CWIC CEOS OGC

Link: <http://commons.esipfed.org/node/4940>

### **Data Stewardship Planning**

Room: Private Dining Room

Session Lead(s): Ruth Duerr, Justin Goldstein

Description: Data stewardship planning session at the winter meeting.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2675>

## **Breakout Sessions, 1:30-3:00**

### **Drupal Working Group: Code Sprint**

Room: Dupont

Session Lead(s): David Bassendine, Adam Shepherd

Description: Since community is such a key component of the success of Drupal, the ESIP Drupal Working Group aspires to build a community of practice that serves the needs of the collective. One of the ways the Drupal community has traditionally approached this has been to sponsor code sprints at meetings, conferences and workshops. These code sprints are central to the effectiveness and draw of the meetups. We are seeking to establish the ESIP Winter Meeting as the official Science on Drupal Code Sprint. Leading up to the Winter meeting, we will be identifying and discussing ideas for topics that we will tackle during this session. This session is open to everyone. If you are having a problem with Drupal, looking to solve a specific issue with Drupal, or are just wanting to peep the ESIP Drupal Working Group in action, please join the conversation at our Drupal.org page at: <https://groups.drupal.org/science-on-drupal> or comment below! VOTE: for your code sprint collaboration area: <https://groups.drupal.org/node/445988>

Session Type: Breakout

Topic tags: Drupal, code, sprint, community, Community of Practice

Link: <http://commons.esipfed.org/node/2713>

### **Birds of a Feather - Free space**

Room: Foggy Bottom

Description: Birds of a feather is free, unscheduled space available if a topic emerges that you'd like to continue a conversation or work space.

Session Type: Other

Link: <http://commons.esipfed.org/node/7763>

### **Digital maturity of federal and federally funded earth sciences - status and next steps**

Room: Mt. Vernon

Session Lead(s): Rick Ziegler, Gary Foley, Lindsay Powers

Description: Please note: This session will be a continuation of sessions 299 and 300 (Jan 7) and will be a working session that broadens the scope of subject matter, and explores publication opportunities - journal articles, white papers, etc. - on matters specific to all 3 sessions. Earth sciences organizations from around the world - including US government agencies, federally funded efforts and academic institutions - have achieved various levels of maturity in taking advantage of our digital age. Concepts of participatory web, software interoperability, technology transfer, scaling/re-use, big data and open science are no longer "new and emerging." They have emerged and - in some cases - are tied to government directives.

Approach: Please note that much of the following may be addressed in sessions 299 and 300, and continued in this session.

Session Type: Breakout

Topic tags: open government, government directives, open science, interoperability, open data, efficiency, transparency, Cyber Infrastructure

Link: <http://commons.esipfed.org/node/7705>

### **(re)Vision 2020 for Earth Science Data Systems**

Room: New Hampshire

Session Lead(s): Chris Lynnes

Description: We're Baaaack! The NASA Vision 2020 Working Group has a candidate set of vision elements. We want to hear from you (again).

Is the Vision:

- too wild or too meh?
- too blurry or too sharp?
- too easy or too hard?

What did we miss? Got ideas for Grand Challenges? Which aspects are applicable to the ESIP Strategic Plan?

Session Type: Breakout

Topic tags: Vision 2020 Discovery

Link: <http://commons.esipfed.org/node/2709>

### **Attribute Convention for Data Discovery: Present and Future**

Room: Potomac

Session Lead(s): Kelly Monteleone

Description: The Documentation Cluster has recently finished revising the Attribute Convention for Data Discovery (see final working draft). Future developments will focus on using groups to provide more discovery details and metadata for other important use cases (access, use, and understanding). An approach to encoding ISO metadata in NcML and HDF will be described for discussion.

Session Type: Breakout

Topic tags: metadata discovery NcML HDF

Link: <http://commons.esipfed.org/node/7754>

## **Dynamic Data Citation**

Room: Private Dining Room

Session Lead(s): Ruth Duerr, Sarah Ramdeen, Andreas Rauber

Description: The primary goal of this workshop will be to explore the feasibility of implementing a data citation model developed by the Research Data Alliance Dynamic Citation Working Group against a number of use cases contributed by members of the ESIP community. These cases focus on dynamic data. The workshop will be run by Andreas Rauber of Vienna University of Technology. Data sets will be provided by ESIP members such as NCAR, Biological and Chemical Oceanography Data Management Office (BCO-DMO), and NSIDC. In addition, a use case about Vector Borne Disease Network EMOD simulation data is being provided by the Hesburgh Libraries of the University of Notre Dame.

In this two session workshop, an overview of the data citation model will be given; followed by brief descriptions of the use case data sets to be examined during the workshop. The remainder of the workshop will focus on assessing the feasibility of applying the data citation model to these datasets, and exploring any issues with the proposed model. The workshop will close with a discussion to determine how to move forward with these steps in combination with the current ESIP Data Citation Guidelines.

A variation of this workshop was previously held in the UK. The attendees included representatives from the UK Natural Environment Research Council data centres, the UK Data Archive of the Economic and Social Research Council, the British Library and DataCite. Through a number of facilitated sessions, the participants of that workshop explored the issues around the proposed model and possible improvements or adaptations for their own user communities; a number of currently used pragmatic solutions were presented and explored, and possible steps forward were proposed for a few of the use cases presented. The report from this workshop can be found on the RDA DCWG website. Simplified summary of the RDA DCWG model goals.

\*\*\*Note the existing ESIP guidelines cover citing subsets but only in a human readable way. This model looks at citing through machine understanding.\*\*\*

1) Citing or making subsets of data citable. For example, if a researcher applies a set of filters, identifying some data which they use in a study. Your organization wants them to cite that specific subset. How do you give it a persistent identifier without having to store a duplicate copy of the data?

2) When you have continuously created data (dynamic or static), or data which is updated, how do you ensure that the citation resolves to the previous version of the data?

Session Type: Breakout

Topic tags: data citation, dynamic data

Link: <http://commons.esipfed.org/node/2677>

## **Breakout Sessions, 3:30-5:00**

### **Birds of a Feather - Free space**

Room: Dupont, Foggy Bottom, Potomac

Description: Birds of a feather is free, unscheduled space available if a topic emerges that you'd like to continue a conversation or work space.

Session Type: Other

Link: <http://commons.esipfed.org/node/7763>

### **Digital maturity of federal and federally funded earth sciences - status and next steps**

Room: Mt. Vernon

Session Lead(s): Rick Ziegler, Gary Foley, Lindsay Powers

Description: See previous session for description.

Session Type: Breakout

Topic tags: open government, government directives, open science, interoperability, open data, efficiency, transparency, Cyber Infrastructure

Link: <http://commons.esipfed.org/node/7705>

## **Earth Science Collaboratory Showcase**

Room: New Hampshire

Session Lead(s): Chris Lynnes

Description: Since the Earth Science Collaboratory was first proposed, a number of technologies and projects have begun to mature to satisfy some of the capabilities that were hypothesized as essential components. This session will be a showcase for those technologies, surveying the capabilities, and possibilities in them. A short discussion will be held at the end to decide whether the ESC Cluster can help the infusion and diffusion of any of the technologies.

Session Type: Breakout

Link: <http://commons.esipfed.org/node/2710>

## **Dynamic Data Citation**

Room: Private Dining Room

Session Lead(s): Ruth Duerr, Sarah Ramdeen, Andreas Rauber

Description: See description from previous session.

Session Type: Breakout

Topic tags: data citation, dynamic data

Link: <http://commons.esipfed.org/node/2677>

## **Poster and Demo Abstracts**

Posters are listed alphabetically, by submitter last name. Poster set-up is during the afternoon break on Tuesday, Jan. 6. Poster presenters are responsible for set-up and teardown of posters after the reception. Foundation staff and volunteers are not responsible for lost posters or poster tubes. \*\* Denotes 2014 FUNding Friday Winner, (D) indicates demo with poster and (T) indicates an ESIP Testbed project.

For all poster and demo abstracts and to comment on the work, please see the ESIP Commons:  
<http://commons.esipfed.org/gallery/Winter%20Meeting%202015>

### **NOAA/NCEP Weather Model Output Data for Decision Support Systems, Including Aggregation of Ensemble Model Output**

Jordan Alpert, Glenn Rutledge

The skill and accuracy of operational atmospheric forecasts from deterministic models have increased, and there are now ensembles of such models that yield a "cloud" of forecasts that provide statistics to improve decisions that protect life, property and commerce. The benefits of improved forecast predictions directly support decision makers for climate, energy, and water users as well as producers. Numerical weather model ensembles are constructed by changing initial conditions to make a cloud of forecasts that attempt to span the space of possible atmospheric realizations which can quantify not only the most likely forecast, but also the uncertainty. However, the ensemble run history has led to an unprecedented increase in data production and information content from higher resolution, multi-model output and secondary calculation products. One difficulty is to obtain the needed subset of information required to estimate the probability of events, and report this information. The calibration required to reliably estimate the probability of events, and honing of decision threshold adjustments (heavy rain, snow, heat waves, weather related traffic, fire weather, etc.) will reduce false alarms for decision makers. To meet the future needs of the ever-broadening user community, ever increasing high volume model outputs, and address these issues on a national and international basis, NOAA implemented the NOAA Operational Model Archive and Distribution System (NOMADS). NOMADS provides real-time and retrospective format independent access to climate, ocean and weather model data and delivers high availability content services as part of NOAA's official real time data dissemination at the NOAA Center for Weather and Climate Prediction web operations center and the archive component at the National Climatic Data Center (NCDC) respectively. An important aspect of the server's abilities is to aggregate the matrix of model output offering access to probability and calibrating information for real time decision making. The aggregation in the content server reports over ensemble component and forecast time in addition to the other data dimensions of vertical layers and position (Latitude/Longitude) for each variable. The unpacking, organization and reading of many binary packed files is accomplished most efficiently on the server side while weather element event probability calculations, incorporating thresholds for more accurate decision support, and display remain for the client (user). Our goal is to reduce uncertainty for variables of interest, e.g. agricultural importance and other discrete events. The weather service operational GFS model ensemble and short range ensemble forecasts can make skillful probability forecasts to alert users if and when their selected weather events will occur. A description of how this framework operates and how it can be implemented using existing NOMADS content services is described.

For all details see: <http://commons.esipfed.org/node/7843>

### **Enriching Earthdata by Improving Content Curation**

Ross Bagwell, Minnie Wong, Kevin Murphy  
NASA Goddard Space Flight Center, CTS, NASA ESDIS, NASA

Since the launch of Earthdata in the later part of 2011, there has been an emphasis on improving the user experience and providing more enriched content to the user, ultimately with the focus to bring the "pixels to the people" or to ensure that a user clicks the fewest amount of times to get to the data, tools, or information which they seek. Earthdata was founded to be a single source of information for Earth Observing System Data and Information System (EOSDIS) components and services as a conglomeration between over 15 different websites. With an increased focus on access to Earth science data, the recognition is now on transforming Earthdata from a static website to one that is a dynamic, data-driven site full of enriched content.

In the near future, Earthdata will have a number of components that will drive the access to the data, such as Earthdata Search, the Common Metadata Repository (CMR), and a redesign of the Earthdata website. The focus on content curation will be to leverage the use of these components to provide an enriched content environment and a better overall user experience, with an emphasis on Earthdata being "powered by EOSDIS" components and services.

Related Collaboration Areas: Information Technology and Interoperability  
For all details see: <http://commons.esipfed.org/node/7845>

### **Discovering Public Data: An Assessment of current metadata practices across the Department of the Interior**

Jonathan Blythe, Jennifer Carlino, Ray Obuch, Lin Zhang  
USGS, DOI OCIO

The Department of the Interior is coordinating an effort for its compliance with the Open Data Policy. All new and updated public data will be part of a federal Public Data Listing at <http://data.gov>, which will be served through the <http://data.doi.gov> catalog. In order to comply with this requirement, the Department conducted a corporate assessment of the metadata practices in order to update its Data Resource Management program. This evaluation started by assessing the current state of metadata practices, by collecting various experiences related to the formalization and codification of metadata procedures in the bureaus and offices of the Department. The evaluation produced four synthetic themes that illustrate the variety of perspectives that may be needed to effectively mobilize metadata from this complex, multi-missioned organization. Actionable outcomes focused on increasing human capital.

Related Collaboration Areas: Decisions, Discovery, Documentation  
Tags: Public Data Listing, Open Data Policy, Evaluation  
For all details see: <http://commons.esipfed.org/node/7826>

### **GISCube, an open source web-based GIS application**

Mazyar Boustani, Chris a. Mattmann, Paul M. Ramirez, Michael J. Turmon  
NASA JPL

Beside some open source GIS libraries and some software like ArcGIS there are comparatively few open source, web-based and easy to use application that are capable of doing GIS processing and visualization. To address this, we present GISCube, an open source web-based GIS application that can store, visualize and process GIS and GeoSpatial data. GISCube is powered by Geothon, an open source python GIS cookbook. Geothon has a variety of Geoprocessing tools such data conversion, processing, spatial analysis and data management tools. GISCube has the capability of supporting a variety of well known GIS data formats in both vector and raster formats, and the system is being expanded to support NASA's and scientific data formats such as netCDF and HDF files. In this talk, we demonstrate how Earth science and other projects can benefit by using GISCube and Geothon, its current goals and our future work in the area.

Tags: GIS, giscube, opensource, NASA, JPL, website  
For all details see: <http://commons.esipfed.org/node/7846>

### **Spatial Analysis Of Sediment Persistence: How Tropical Cyclone Events Affect Sedimentation In The Gulf Of Mexico**

Kaylin Bugbee, Sara Graves, Sandy Ebersole  
University of Alabama in Huntsville, Geological Survey of Alabama

Tropical storms and hurricanes are a common threat along the Gulf of Mexico coast. In the ten year period from 2000 to 2010, Louisiana, Mississippi, Alabama and the Florida panhandle were affected by 28 tropical storms, seven of which were hurricanes. Tropical cyclones have many impacts on the region including increased rainfall and flooding. These resulting flood waters carry large amounts of suspended sediments. These suspended sediments often migrate to the Gulf of Mexico coastal areas and are a factor in aquatic habitat health, water quality, conservation efforts, and commercial endeavors. This study assesses Gulf Coast areas most often affected by suspended sediments following tropical cyclones by utilizing the Sediment Analysis Network for Decision Support (SANDS) dataset, created from MODIS and Landsat imagery and provided by the Global Hydrology Resource

Center. The results of this study will be used to map areas in the Gulf of Mexico where suspended sediments consistently occur following tropical cyclones in order to aid in decision-making processes.

Related Collaboration Areas: Geospatial, Student Cluster

Tags: Hurricane, water resources, sediment disturbance, geospatial, Landsat, MODIS, Decision Support

For all details see: <http://commons.esipfed.org/node/7780>

### **Big Data for a Big Ocean - Preserve, Discover, Access, and Use**

Kenneth Casey, John Relph, Yuanjie Li

NOAA National Oceanographic Data Center, NOAA National Oceanographic Data Center (NODC)

Goals: Tiers of Stewardship

The mission of the NOAA National Oceanographic Data Center (NODC) is to provide scientific stewardship of marine data and information. NODC provides various levels of stewardship of our data; long term preservation and basic access to data is the absolute minimum level of stewardship. NODC engages its partners around the globe to provide additional levels of stewardship.

Challenges: The "Big Ocean"

With increasing numbers of ocean observing systems, including automated ship-based systems, autonomous underwater and airborne vehicles, satellite observations, and other in situ systems, combined with numerical models, NODC has clearly entered the age of Big Data. The challenges of stewarding this growing variety, volume, and velocity of incoming "Big Ocean" data, while also maintaining the veracity of those data, will continue to grow.

Strategy: Automation and Standardization

NODC continues to adapt to meet these challenges by taking numerous approaches, including automating the acquisition and ingest of incoming data; deploying standardized, machine-consumable data discovery services; and providing interoperable data access, visualization, and subset mechanisms. NODC also supports the community of data producers by developing tools to help create more machine-ready ocean observation data and information. NODC undertakes projects to investigate how, for example, to more effectively incorporate cloud storage, access, and processing services into existing workflows and systems.

Related Collaboration Areas: Data Preservation, Discovery, Documentation, Visualization

Tags: NODC, Big Data, Oceanographic, data stewardship, Discovery, Data access, data ingest, automation

For all details see: <http://commons.esipfed.org/node/7775>

### **Rule-Based Curation in the DataNet Federation Consortium: Policies to Policies Using iRODS**

W. Christopher Lenhardt

RENCI, DICE, DataNet Federation Consortium, School of Information and Library Science UNC-CH

This poster will present efforts to translate human-readable data curation policies into rule-based curation through middleware (iRODS). As part of the DataNet Federation Consortium, the interaction of ISO 16363, basic digital data curation and data lifecycle management, information technology security, and end-user considerations are being examined by the DFC Policies, Facilities and Operations, and Technology Groups. The goal is to identify overlap and synergies in the human-readable policies and to map those onto iRODS microservices. Activities such as this will have a longer-term impact in areas such as implementing funder requirements for such things as data management plans, and to help ensure data preservation and facilitate reuse throughout the data lifecycle.

Related Collaboration Areas: Data Preservation

For all details see: <http://commons.esipfed.org/node/7809>

### **Data System for HS3 airborne field campaign**

Helen Conover, Manil Maskey, Rahul Ramachandran, Sara Graves, Kaylin Bugbee, Michael McEniry

ESO, University of Alabama in Huntsville, NASA/MSFC

Hurricane and Severe Storm Sentinel (HS3) is a NASA airborne field campaign aimed at better understanding the physical processes that control hurricane intensity change. HS3 will help answer questions related to the role of environmental conditions and internal storm structures to storm intensification. Due to the nature of the questions that HS3 mission is addressing, it involves a variety of in-situ, satellite observations, airborne data, meteorological

analyses, and simulation data. Variety of datasets present various data management challenges for HS3. The methods used for airborne data management differs greatly from the methods used for space-borne data. In particular, metadata extraction, spatial and temporal indexing, large number of instruments and subsequent variables are few data management challenges unique to airborne missions. A robust data system is required to successfully help HS3 scientist achieve their mission goals. Furthermore, the data system also needs a data management that assists in broader use of HS3 data to enable future research activities. The Global Hydrology Resource Center (GHRC) is considering all these needs and designing a data system for HS3. Experiences with past airborne field campaign puts GHRC in a good position to address HS3 needs. However, the scale of the mission along with science requirements separates HS3 from previous field campaigns. HS3 data system will include automated services for geo-location, metadata extraction, discovery, and distribution for all HS3 data. To answer the science questions, the data system will include a visual data exploration tool that is fully integrated into the data catalog. The tool will allow visually augmenting airborne data with analyses and simulations. Satellite data will provide contextual information during such explorations. All HS3 tools will be supported by an enterprise service architecture that will allow scaling, easy integration of new tools and existing services, and integration of new ESDIS metadata and security guidelines.

Related Collaboration Areas: Information Technology and Interoperability, Science Software, Visualization

Tags: data systems, field campaigns

For all details see: <http://commons.esipfed.org/node/7789>

### **Unidata: Helping the University Education and Research Community Access and Use Real-time Weather Data**

Ethan Davis, Josh Young, Mohan Ramamurthy, Doug Dirks  
UCAR/Unidata, UCAR Unidata

Unidata was formed nearly 30 years ago at the request of the university atmospheric science community (and funded by the NSF) to facilitate access to a broad range of observation and forecast data (e.g., station, satellite, radar, and model) and to support the community's use of that data.

Unidata works with the broad Earth Science community, including agencies like NOAA's National Weather Service and other data providers, to gain access to important datasets and make them widely available. Data are made available to the community in two main ways. First, the IDD/LDM system pushes data to over 200 institutions around the world in near real-time (generally with sub-second latencies, depending on a sites network capacity) for local access. Second, Unidata works to make sure data are available for remote access through various web services such as OPeNDAP and OGC WMS.

Unidata provides and supports software tools for managing, analyzing, and visualizing data. For managing data, Unidata provides and supports a number of software packages including netCDF, LDM, and THREDDS Data Server (TDS). For analysis and visualization, Unidata provides and supports the IDV, GEMPAK, McIDAS-X, and AWIPS II. Unidata developed many of these software tools, but also provides support for a number of externally developed software tools (e.g., AWIPS II, GEMPAK, and McIDAS). All software developed at Unidata is free and open source software (FOSS).

Besides providing software development and support, Unidata works to engage with geoscience educators and researchers by offering training, hosting workshops, and serving as their advocate in securing open data access.

For all details see: <http://commons.esipfed.org/node/7748>

### **Enabling Ongoing Access to Data Products and Services When Dependencies are No Longer Supported**

Robert R. Downs  
Columbia University

Online scientific data and related products and services can be developed in a distributed manner across systems and organizations, creating dependencies that rely on continuing operations of coordinated services over time. Dependencies can include other online products and services that rely on disparate chains of hardware and software maintained by separate organizations with independent operational capabilities, support, and sustainability models. In the future, some aspects of such distributed networks of dependent capabilities may fail, requiring mitigation to maintain the online capabilities that users expect. Case studies are offered to describe opportunities for

enabling continuing access to scientific data products and services that are dependent on external systems that may no longer be viable.

Related Collaboration Areas: Data Preservation, Preservation and Stewardship  
Tags: Science Data Products, Coordinated Service Networks, online data, Contingency Planning  
For all details see: <http://commons.esipfed.org/node/7787>

### **ESDIS Standards Office Information**

Allan Doyle, Yonsook Enloe, Helen Conover  
ESO

We present an overview of the ESDIS Standards Office (ESO). Information is provided about the ESO including an overview of the standards currently approved for use in NASA Earth Science Data Systems as well as standards and specifications currently under development. We also provide an overview of the ESO Standards Process. The ESO is also assisting the Common Metadata Repository (CMR) Systems Engineering Team with a series of reviews of metadata documents. Information about this is also provided. This poster is meant to provide an introduction to those unfamiliar with the ESO as well as to provide information for discussion with ESO Staff members who will be presenting the poster.

Related Collaboration Areas: Information Technology and Interoperability  
Tags: ESDIS, ESO, Standards, CMR, datasystems  
For all details see: <http://commons.esipfed.org/node/7781>

### **Implementing Ecological Metadata Language in a Cross-Disciplinary Environmental Monitoring Database**

James Duncan  
Vermont Monitoring Cooperative

The Vermont Monitoring Cooperative, a collaboration of the University of Vermont's Rubenstein School of Environment and Natural Resources, the US Forest Service and the Vermont Agency of Natural Resources, has been collecting and federating a range of data about Vermont's forested ecosystems since the early 1990s. As part of its central mission, the VMC maintains a database ranging from near-real-time air-borne mercury sampling to avian demographic surveys to forest health measurements to long-term soil monitoring data, among others. In addition to the data, a documentation system developed over the years has driven a website to expose metadata and, in most cases, data downloads, but a need for more standardization in this metadata documentation led to the implementation of Ecological Metadata Language (EML). EML was developed for use in the ecological sciences, necessitating a fairly flexible standard to deal with the wide range of potential datasets, and is the standard employed by the Long-Term Ecological Research Network in the United States. This poster describes the process of upgrading VMC's data documentation to EML, the gains realized in delivering discoverable data via VMC's metadata-driven website, future plans for federation to other catalogs, and the resulting status of the VMC database and website, which can be used to discover millions of records of ecological monitoring information in the state.

Tags: metadata, documentation, Ecological informatics  
For all details see: <http://commons.esipfed.org/node/7750>

### **ToolMatch Use Case Extension\*\***

Matt Ferritto, RPI

The ToolMatch service was developed with the intent to provide data users with the means to match their data collections with a comprehensive list of useful, appropriate tools, and to provide data tool developers with data collections that will work with their tools. As such ToolMatch had an initial scope of two use cases, the first of which was the semantic matching of data collections with tools. This would allow data users to find and choose among a list of otherwise separate and potentially hard to find tools that could work with their data collections. The second (and more difficult) of these use cases was the converse: given a tool, semantically find what data collections that the tool can use. If the first use case is analogous to having nails and looking for a hammer, then the second use case can be compared to having a hammer and looking for nails. It is much more difficult to find data collections that may work with a given tool, since a tool user might not necessarily know what to look for. Using the ToolMatch service, a tool user could easily find a data collection to use with their tool. In both of these use cases, wasted time and effort searching for the correct tool or data collection can be reduced or avoided completely. The focus of this

thesis will be on the implementation of these two use cases, as well as an extension of the first use case, where a data user with certain semantics for a given data collection (such as a domain model) can find tools that can be used with the content of that data. This is an important issue due to the fact that a certain data collection content may not be appropriate for a tool within a certain domain model. For example, rainfall or topographic data content that is part of a larger Hydrological model can be matched to tools that the model as a whole might not be able to match. This expands the scope of the initial use case in that data collection content requires stricter matching than just the characteristics of data collection. The requirements of this use case involve modification and expansion to the ToolMatch conceptual model and ontology to allow for semantic matching between data content and tools. These changes will also be reflected in the ToolMatch web service, which allows users to make add, update, or delete instances of the ToolMatch ontology without having to have a full understanding of ontologies.

Related Collaboration Areas: Semantic Web

Tags: semantic matching

For all details see: <http://commons.esipfed.org/node/7735>

### **Towards Executable Provenance Graphs for Reported Results in Research Publications\*\***

Linyun Fu, Xiaogang Ma, Stace Beaulieu, Patrick West, Peter Fox  
Rensselaer Polytechnic Institute, Woods Hole Oceanographic Institution, TWC

Results in research publications often are not replicable because the publications and results described in the publications are separated from the underlying collection and analysis of data. Using general provenance ontologies such as PROV-O, the new W3C standard adopted in 2013, proves to be an effective way to keep track of the lineage of the source data and the changing processes leading to the final results.

This poster presents how we are extending PROV-O to create a specialized provenance model for the replication of the process from data transformation to reporting of results in research publications. The proposed model can even be used to validate the scientific conclusions by allowing readers to adapt existing experiments reported in the papers and carry out their own studies.

Related Collaboration Areas: Preservation and Stewardship, Semantic Web

Tags: Provenance; Ontology

For all details see: <http://commons.esipfed.org/node/7808>

### **Event-Driven Cyberinfrastructure Technologies Supporting The Disaster Lifecycle**

Sara Graves, Ken Keiser, Helen Conover, Rahul Ramachandran, Manil Maskey, Ajinkya Kulkarni  
University of Alabama in Huntsville, NASA/MSFC

Event-driven data delivery (ED3) is being developed to provide data preparedness through the automation data access and processing for decision support as the result of disaster events. Through a NASA Applied Science award, the Information Technology and Systems Center and Atmospheric Science Department, both at the University of Alabama in Huntsville (UAH), and the Geological Survey of Alabama, are collaborating to refine and apply the ED3 functionality for data response to disasters. The team is working closely with decision support systems, such as the SERVIR project and state and regional agencies and programs in Alabama, to help define the capabilities that will best suit the end users for a variety of disaster events. This poster depicts Event Listener and Subscription Management components of ED3 and shows how they can be used to drive the generation of Data Albums for disaster event occurrences. The ED3 project will demonstrate the reusability of this technology for other disaster events and in support of multiple decision support systems and end users.

Related Collaboration Areas: Decisions, Science Software

Tags: Disaster, Disaster Response, Event-Driven

For all details see: <http://commons.esipfed.org/node/7788>

### **Science on Drupal**

Sara Graves, Helen Conover, Ajinkya Kulkarni, Manil Maskey, Bruce Caron, Rahul Ramachandran  
University of Alabama in Huntsville, EDDC, NASA/MSFC

The "NASA Science on Drupal Central (NSODC)" ACCESS project has been working hard to build Drupal capabilities to support NASA Earth science activities. The project has been developing key new open-access modules for Earth science data- and metadata-rich websites. NSODC is also bringing together NASA Earth

science Drupal web developers into a community of purpose, where they can share code, knowledge, lessons learned, and achievements. All software created by this project will be distributed as open-source code. The project's ScienceOnDrupal twitter and Google+ accounts harvest Drupal insights from across the web. The central website: scienceondrupal.org is a hub for using Drupal in the sciences. The project has led and participated in Science on Drupal sessions at various conferences and is active in the ESIP Federation Drupal Working Group.

Related Collaboration Areas: Drupal Working Group

Tags: Drupal, Earth Science Data, open-source software

For all details see: <http://commons.esipfed.org/node/7721>

## **Content Model Use and Development to Redeem Thin Section Records**

Denise J. Hills

The National Geothermal Data System (NGDS) is a catalog of documents and datasets that provide information about geothermal resources located primarily within the United States. It has made large quantities of geothermal-relevant geoscience data available to the public by creating a national, sustainable, distributed, and interoperable network of data providers. NGDS facilitates interoperability through the use of content models (CMs). CMs provide a schema (structure) for submitted data. Schemas dictate where and how data should be entered. CMs use templates that simplify data formatting to expedite use by data providers. The NGDS uses CMs utilizing USGIN (U.S. Geoscience Information Network) schema. USGIN extends beyond geothermal data to all geoscience data.

Countless thin sections have been taken from oil and gas well cores housed at the Geological Survey of Alabama (GSA), and many of those thin sections have related photomicrographs. Record-keeping for these thin sections has been scattered at best, and it is critical to capture their metadata while the content creators are still available. The GSA, using USGIN CMs, has tested and refined a CM for thin sections and thin section photos. This allows GSA to register the thin sections and related samples with SESAR (System for Earth Sample Registration) and have IGSNs (International Geo Sample Number) assigned. By developing a process model that provides the necessary checks and balances to assure quality of information, the GSA is able to make these previously dark data to light.

Related Collaboration Areas: Data Preservation

For all details see: <http://commons.esipfed.org/node/7722>

## **What's Up these Days with Persistent Identifiers for Earth Science Data? (T)**

Nancy Hoebelheinrich, Greg Janee

Knowledge Motifs, University of California at Santa Barbara

It has been over a year since the ESIP Data Stewardship Committee received the final report on the findings of the testbed implementation of selected identifier schemes

([http://wiki.esipfed.org/index.php/Executive\\_Summary\\_of\\_Results\\_from\\_Report\\_by\\_Hoebelheinrich\\_and\\_Janee](http://wiki.esipfed.org/index.php/Executive_Summary_of_Results_from_Report_by_Hoebelheinrich_and_Janee)),

and over 3 years since the publication of the article providing an assessment of those identifier schemes with recommendations for an implementation testbed (<http://link.springer.com/article/10.1007%2Fs12145-011-0083-6>).

What has been happening with the assignment of persistent identifiers for data since then? How many data producers are assigning persistent identifiers to datasets or data collections and to what kinds of organizations do they belong? What are the issues that have arisen as organizations develop procedures and practices to incorporate the assignment and use of persistent identifiers?

This poster will summarize the issues included in the final report, but also begin to address and seek input on some of the research issues associated with the real-time, production level implementation of persistent identifiers by Earth Science data producers, such as:

- In the last 3 - 4 years, how many institutions have begun assigning DOIs / ARKs to "research data"
- What type of institution (governmental agency, academic department, non-profit research organization, etc.) is assigning persistent identifiers
- For what purpose are identifiers being assigned (citation, inventory, etc.)
- What are the advantages / disadvantages of assigning "opaque" identifiers versus those with inherent semantics
- At what level of granularity (data collection, dataset, singular data file., etc., )
- What is the relationship b/w / among identified items (data collection ? dataset, versions, etc.,)
- Who "owns" the identifiers, and what are the issues associated with transferring ownership

Related Collaboration Areas: Data Preservation  
Tags: Persistent identifiers, DOI, data stewardship  
For all details see: <http://commons.esipfed.org/node/7794>

### **Extending the ToolMatch Service by Expanding Community Engagement (T)**

Nancy Hoebelheinrich, Christopher Lynnes, Patrick West, Matt Ferritto  
Knowledge Motifs, NASA Goddard Space Flight Center, Rensselaer Polytechnic Institute

In order to make further progress on the viability and robustness of the ToolMatch service, much more instance data is needed to add to the knowledge store in the form of different kinds of visualization tools, and many more datasets from a variety of domains. Further populating the knowledge store will not only confirm that the ToolMatch service meets the two initial use cases, and the third, new use case, but also expand the applicability of the service to other domains within the ESIP Federation. As we have begun adding information to the underlying Knowledge Store, it has become increasingly clear that employing other sources of information such as data catalogs, and data or tool registries is not only advisable but critical to leverage and scale a service like ToolMatch. In-depth analysis of the types of datasets, visualization tools, and technologies used by these data catalogs and registries will be necessary, however, in order to understand how the ToolMatch service can use them in a practicable, scalable manner. This kind of analysis will also help the ToolMatch team move toward the goal of demonstrating how the service can be incorporated into existing sets of information services found at data and archive centers. Short term outcomes for the ToolMatch service supported by funding from the ESIP Products & Service Committee's Testbed are:

- Address the feasibility of integrating Semantic web based services with other information services such as data / tool / service catalogs and registries in support of data user and data tool developer information needs;
- Investigate the applicability of the ToolMatch service to other domains within the ESIP Federation, i.e., USGS and NOAA;
- Incorporate some relevant datasets and visualization tools from the other domains mentioned above into ToolMatch, if applicable;
- Identify factors that would influence the successful incorporation of a ToolMatch service into at least one data catalog;
- Identify factors that would influence the successful incorporation of a ToolMatch service into the services offered by an established data center such as that of NASA's Goddard Space Flight Center or Jet Propulsion Laboratory.

Related Collaboration Areas: Products and Services, Semantic Web  
Tags: semantic matching, ToolMatch, Semantic Web technologies  
For all details see: <http://commons.esipfed.org/node/7757>

### **Human & Machine Actionable Data Citations**

Joseph A. Hourcle, Ruth Duerr, Robert Downs  
Goddard Space Flight Center, NSIDC, Data Conservancy, CIESIN

Earlier this year the Joint Declaration for Data Citation Principles was released. The preamble of the declaration states "These principles recognize the dual necessity of creating citation practices that are both human understandable and machine-actionable." A group was formed to discuss the issues regarding implementing these principles and currently has a paper in process, "Achieving human and machine accessibility of cited data in scholarly publications" with a proposal of how to achieve the goal in the preamble. We will present an short overview of both the data citation principles and our proposal and solicit comments on the paper during its review period.

Related Collaboration Areas: Discovery, Documentation, Information Technology and Interoperability, Preservation and Stewardship  
Tags: data citation  
For all details see: <http://commons.esipfed.org/node/7768>

## **Digital Object Identifiers (DOIs) Usage and Adoption in US Geological Survey (USGS)**

Viv Hutchison, Lisa Zolly, Mike Frame, Giri Palanisamy  
US Geological Survey, Oak Ridge National Laboratory

Addressing grand environmental science challenges requires unprecedented access to easily understood data that cross the breadth of temporal, spatial, and thematic scales. From a scientist's perspective, the big challenges lie in discovering the relevant data, dealing with extreme data heterogeneity, large data volumes, and converting data to information and knowledge. Historical linkages between derived products, such as publications and associated datasets, have not existed in the earth science community. The US Geological Survey's Core Science Analytics Synthesis & Libraries, in collaboration with Department of Energy's Oak Ridge National Laboratory (ORNL) Mercury Consortium (funded by NASA, USGS and DOE), established a Digital Object Identifier (DOI) service for USGS data, metadata, and other media. This service is offered in partnership through the University of California Digital Library EZID service. USGS scientists, data managers, and other professionals can generate globally unique, persistent and resolvable identifiers for any kind of digital objects. Additional efforts to assign DOIs to historical data and publications are also underway. New policies will require DOIs in metadata records such that the data used in the research can be better identified. USGS is using DOI identifiers to cite data in journal articles, web-accessible datasets, and other media for distribution, integration, and in support of improved data management practices. The poster will discuss the current DOI efforts within USGS, including adoption, challenges, and future efforts necessary to improve access, reuse, sharing, and discoverability of USGS data and information.

Related Collaboration Areas:

Tags: DOI, persistent identifier, digital object identifier

For all details see: <http://commons.esipfed.org/node/7718>

## **Big Data Challenges in a Data Center Workflow**

Eric Kihn  
NGDC

The Mission of NOAA's National Geophysical Data Center (NGDC) is to provide long-term scientific data stewardship for the Nation's geophysical data, ensuring quality, integrity, and accessibility. NGDC provides stewardship, products, and services for geophysical data from our Sun to Earth and Earth's sea floor and solid earth environment, including Earth observations from space. As part of its mission, NGDC executes preservation workflows which include ingest, quality control, metadata generation, product generation, and development of access methods for diverse data types. Each phase of proper stewardship involves challenges when it comes to Big Data. This poster will look at Big Data as it interacts with, and is supported by, the stewardship workflow of a National Data Center. We will present tools and techniques as well as identify remaining challenges as we continue to march into the Big Data era.

For all details see: <http://commons.esipfed.org/node/7790>

## **Global Precipitation Measurement (GPM) Mission Applications (D)**

Dalia Kirschbaum  
Goddard Space Flight Center, NASA

The Global Precipitation Measurement (GPM) Mission was launched in February, 2014 and uses advanced space-borne instruments to measure global precipitation. Through improved measurements of rain and snow, precipitation data from GPM has revealed new information on hurricane eyewalls and intensity, measured hazard-triggering rainfall events, provided inputs into climate and land surface models, and offers new insights into agricultural productivity and world health. GPM data enables a diverse range of applications across agencies, research institutions and the global community. This poster will provide some highlights of GPM data, present innovative new applications and provide an overview of data access and end user engagement opportunities.

Related Collaboration Areas: Data Management Training, Data Preservation, Earth Science Collaboratory, Energy and Climate, Products and Services

For all details see: <http://commons.esipfed.org/node/7782>

## **HDFCRAFT: Making Data Fun! (D) \*\***

Hyo-Kyung Lee

## The HDF Group

Minecraft is a popular game among K-12 students. Minecraft allows players to manipulate textured cubes in a 3D generated world. This poster presents how to use NASA Earth science data in HDF to create a Minecraft world. The goal is to give students early exposure to NASA Earth data and let them explore a world observed by NASA satellites with Minecraft.

Related Collaboration Areas: Visualization

Tags: minecraft, HDF, HDF-EOS

For all details see: <http://commons.esipfed.org/node/7742>

## Augmenting Basic Web Services with Middleware Services and Interfaces

David Meyer, Jason Werpy, Robert Quenzer

US Geological Survey, LP DAAC (Information Dynamics), LP DAAC (SGT)

Over the past few years many Data Providers have implemented services that allow for web based (HTTP) interfaces to manipulate, organize, modify, and deliver Earth Science Data. This web architecture provides the foundation for streamlining of Earth Science Users utilization of and interaction with the Data. However, critical components are missing and need to be developed in order to increase the capabilities, potential, and reach of these services. Middleware services represent a class of Data Services that are able to communicate their capabilities more clearly and effectively with Science Data Users while also leveraging the more raw web services on the back end. A Middleware layer of a services architecture functions to coordinate the interactions of the users with the core web services. This simplifies execution, parameter selection, data integration, data delivery, and data analysis activities. This presentation will outline how the LP DAAC has utilized core services to provide basic access to data, data manipulation, and processing and then followed those efforts with the development and implementation of Middleware on top of those core services to augment capabilities and create workflows necessary to enable Science Users to perform meaningful science activities and analysis faster than before. The Middleware layer acts as the "glue" that allows all these separate services to work together. By moving the algorithms that process and organize data closer to the Data Archive and enabling access to them via web services fronted by Middleware services, the LP DAAC helps Science users to do better, less expensive, and more expansive science much faster than they ever could before.

Related Collaboration Areas: Information Technology and Interoperability

Tags: OPeNDAP Middleware Interface

For all details see: <http://commons.esipfed.org/node/7848>

## An Information Architect's View of Earth Observations for Disaster Risk Management

Karen Moe, John Evans, Stuart Frye, Daniel Mandl, Pat Cappelaere, Kevin Dobbs

Satellite observations play a significant role in supporting disaster response and risk management, however data complexity is a barrier to broader use especially by the public. In December 2013 the Committee on Earth Observation Satellites Working Group on Information Systems and Services documented a high-level reference model for the use of Earth observation satellites and associated products to support disaster risk management within the Global Earth Observation System of Systems context. The enterprise architecture identified the important role of user access to all key functions supporting situational awareness and decision-making. We focus on the need to develop actionable information products from these Earth observations to simplify the discovery, access and use of tailored products. The information architecture can address usability challenges to transform sensor data into actionable information, based on the terminology of the emergency management community responsible for informing the public. This poster describes an approach within ESIP to collecting relevant material from the disasters and risk management community to address the end user needs for information. The resulting information architecture is needed to address the structural design of the shared information in the disasters and risk management enterprise. Key challenges are organizing and labeling information to support both online user communities and machine-to-machine processing for automated product generation.

The ESIP Disaster Lifecycle Cluster has been established to facilitate connections and coordinate efforts among data providers, managers and developers of disaster response systems and tools, and end-user communities within ESIP. The cluster provides a forum to share insights and best practices on user needs, common procedures, common barriers and to translate those needs into system architecture practices. The cluster intends to explore emerging technologies and match to needs to bridge the gap between developers and users. A major contribution is

the ESIP sponsored Collaborative Common Operating Picture (C-COP) Testbed, which helps participants identify and test ESIP member data sets to be recognized as trusted data sources for agencies and organizations responding to disasters.

Related Collaboration Areas: Disaster Life Cycle  
Tags: Disasters Risk Management, Actionable Information  
For all details see: <http://commons.esipfed.org/node/7793>

### **Converting mb-system Files in Windows and Apple \*\***

Kelly Monteleone, John Relph, Jason Ninneman, Rob Carver  
TERA, a CH2M Hill Company, NOAA National Oceanographic Data Center (NODC)

This poster presented the results of the Summer 2014 FUNDing Friday proposal to create a GUI for MB-system. Mb-system is a unix based program for multibeam sonar (Bathymetry) data. Mb-system was supported by NOAA from 2002 to 2004. It does not work in windows and is complicated to install and use. The goal was to create a one click installer that will install Cygwin, mb-system, and supporting files on a windows system. And pull out a python wrapper for Apple and non-cygwin installation. This will provide a GUI to export the data in xyz and asii.

Related Collaboration Areas: Geospatial, Information Technology and Interoperability, Products and Services, Science Software  
For all details see: <http://commons.esipfed.org/node/7760>

### **Architectures Toward Reusable Science Data Systems**

John F Moses

Science Data Systems (SDS) comprise an important class of data processing systems that support product generation from remote sensors and in-situ observations. These systems enable research into new science data products, replication of experiments and verification of results. NASA has been building systems for satellite data processing since the first Earth observing satellites launched and is continuing development of systems to support NASA science research and NOAA's Earth observing satellite operations. The basic data processing workflows and scenarios continue to be valid for simple remote sensor observation research projects as well as for the complex multi-instrument operational satellite data systems being built today. System functions such as ingest, product generation and distribution need to be configured and performed in a consistent and repeatable way with an emphasis on scalability. This paper will examine the key architectural frameworks of two satellite data processing systems currently in operation and under development that make them suitable for scaling and reuse. By highlighting key elements and implementation experience we expect to find architectural components that will outlast their original application and be readily adaptable and reusable. Concepts and principles are explored for guiding SDS developers and strategists.

Related Collaboration Areas: Information Technology and Interoperability  
For all details see: <http://commons.esipfed.org/node/7821>

### **Enhancing and Educating with the WxSat Mobile App \*\***

Kyle Nelson, Margaret Mooney

Satellite imagery is becoming more prevalent in our daily lives. With ease of access online and its display increasingly often in local and national TV weather segments, interest is growing. If we are to become a more Weather Ready Nation, it is critical to enhance the weather literacy of the general public, weather enthusiasts and decision-makers worldwide who view satellite and other earth data products and use them to make informed decisions. For this project, the free WxSat mobile app is modified and enhanced to include the full product suite available in the Space Science and Engineering Center's (SSEC) RealEarth web interface, and new features such as the ability to layer multiple products, save favorites with custom view locations and times, view information about individual products and more. By expanding the available datasets beyond the original Visible, Infrared and Water Vapor satellite imagery, WxSat becomes a more powerful and robust real-time weather analysis and case study tool for the public, scientists and decision makers. The curation of 25 mobile-friendly product information web pages for the most frequently accessed datasets establishes the WxSat app and SSEC RealEarth as educational earth science data viewers. The information pages provide users with a qualitative explanation of the product being viewed, empowering them to more accurately interpret what they see on the screen and infer what effects may be experienced at ground level. Product descriptions are written using plain language so that readers will not be

bombarded by jargon and acronyms commonly used in the atmospheric sciences. Instead, they will be graced with a scientifically sound description that is readily applied to the dataset they are viewing. Product pages for visible, infrared and water vapor imagery include annotated images to provide a visual conceptual model for the user to apply when viewing these satellite imagery products in real time via WxSat, SSEC RealEarth or another source. The enhanced WxSat app will make earth science data more readily available to everyone via mobile devices, allowing them to have realtime weather information in the palm of their hand. Carefully crafted product information pages will aid in building a more Weather Ready Nation through enhancing the weather literacy of the general public, weather enthusiasts and decision-makers worldwide who view satellite and other earth data products via WxSat and SSEC RealEarth

Related Collaboration Areas: Climate Education Working Group  
For all details see: <http://commons.esipfed.org/node/7860>

### **Expanding OceanLink using Facebook crowdsourcing**

Sean Nguyen  
EarthCube

This poster will explain what the OceanLink Project does and what the plans are for Sean Nguyen (Undergraduate Research Assistant) to expand on the current implementation. Today's scientific investigations are producing large numbers of scholarly products. These products continue to increase in diversity and complexity as researchers recognize that scholarly achievements are not only published articles but also datasets, software, and associated supporting materials. OceanLink is an online platform that addresses scholarly discovery and collaboration in the ocean sciences. The OceanLink project leverages Semantic Web technologies, web mining, and crowdsourcing to identify links between data centers, digital repositories, and professional societies to enhance discovery, enable collaboration, and begin to assess research contribution. Currently OceanLink is creating links between the data sources to better connect the data. Sometimes the data mining system will come up with incorrect links. These links need to be manually fixed by a man in the loop. Sean will be creating an application that searches social media for individuals who have the qualifications to identify incorrect links and possibly help create parameters for new links to be formed.

Related Collaboration Areas: Information Technology and Interoperability  
Tags: crowdsource, OceanLink, Geosciences  
For all details see: <http://commons.esipfed.org/node/7784>

### **Integrating and Visualizing Sea Ice Charts and Cruise Tracks using Linked Open Data and Open Source Tools (D)**

Evan W. Patton, Ruth Duerr, Peter Fox, Siri Jodha S. Khalsa, Deborah L. McGuinness, Peter Pulsifer, Cassidy Thompson, Rui Yan  
Rensselaer Polytechnic Institute, NSIDC, Data Conservancy, University of Colorado, CIRES

The Semantic Sea Ice Interoperability Initiative (SSIII) is an NSF-funded research project aimed at making sea ice data more useful to more people using semantic technologies, and includes domain and science data experts at the National Snow and Ice Data Center (NSIDC) and knowledge representation experts from Rensselaer Polytechnic Institute (RPI). SSIII has developed ontologies for modeling sea ice charts as open linked data, making those data available on the web to a wider audience of researchers using web-standard representation and query languages. We will demonstrate a prototype web interface that incorporates the linked data published by SSIII with cruise tracks and dataset metadata published by Woods Hole Oceanographic Institute's Biological and Chemical Oceanography Data Management Office (BCO-DMO). We will show how one can use this interface to (1) determine sea ice concentration, form, and stage of development in the region around a particular cruise during a specified time frame; (2) animate sea ice changes during a given time frame; and (3) highlight links to relevant datasets based on user selections and metadata present in linked data repositories. This interface is built on an open source software stack and may serve as an exemplar to other projects looking to model and integrate geospatial and temporal datasets distributed across the Web.

Related Collaboration Areas: Semantic Web  
Tags: open-source software, Linked Open Data, geospatial portal, Semantic Web technologies  
For all details see: <http://commons.esipfed.org/node/7807>

## **Ontology Alignment with Mappings published in the Purdue Research Repository (T)**

Line Pouchard

A current focus of the Semantic Web Cluster is ontology reuse and ontology mapping. Currently, a new ontology tends to be created for each new project, resulting in duplication of work and lack of interoperability. In this project we are testing the ESIP ontologies by evaluating their potential for ontology mapping. In particular, we are providing mappings between the Semantic Web for Earth and Environmental Terminology (SWEET) and ENVO, the Environmental Ontology by using Agreement Maker Light, an ontology alignment algorithm that won first place in the Ontology Alignment Initiative 2014 campaign. We are also providing mappings between SWEET and the ontologies from the Semantic Sea Ice Interoperability Initiative. The mappings in the form of rdf triples are published in the Purdue University Research Repository and given a DOI. Publication in the repository guarantees open access and preservation of the mappings.

Related Collaboration Areas: Semantic Web

Tags: ontologies, ontology mappings

For all details see: <http://commons.esipfed.org/node/7772>

## **NASA's Implementation of the President's Climate Data Initiative**

Ana Privette, Curt Tilmes, Rahul Ramachandran, Kaylin Bugbee, Jason Duley, Fabien Laurier

In June 2013 President Obama launched the Climate Action Plan to cut carbon pollution, prepare communities for the impacts of climate change, and lead international efforts to address this global challenge. The plan recognizes that even as we act to curb the carbon pollution that is driving climate change, we must also prepare our citizens and communities for the climate impacts that are already underway across the country. One of the efforts described in that Climate Action Plan is the Climate Data Initiative, a broad effort to leverage the federal government's extensive, freely-available climate-relevant data resources to spur innovation and private-sector entrepreneurship in order to advance awareness of and preparedness for the impacts of climate change. NASA was asked to implement this initiative by working closely with several other Federal agencies in order to identify key data and resources and make those available to the data innovators community. The CDI was launched in March 2014 and leverages commitments from government and the private sector to unleash data and make it accessible to the larger community. It builds on the White House's other Open Data Initiatives--in areas such as health, education, and safety. The Climate Data Initiative is organized around thematic areas of climate-change risk and impact, which are closely aligned with the findings of the Third U.S. National Climate Assessment. A team of federal experts works in each thematic area to build and deploy resources on Climate.Data.gov for use by data innovators. Several themes have been released to date including Coastal Flooding, Food Resilience, Water and Ecosystem Vulnerability. Going forward, the Initiative will expand to include data from other themes areas including Energy, Human Health, and Transportation.

For all details see: <http://commons.esipfed.org/node/7797>

## **The legacy of the Bering Sea Project: archival and preservation of the project data for current and future research**

Don Stott, James Moore, Steve Williams, Amanda Orin

The Bering Sea Integrated Ecosystem Research Program (BSIERP) developed and supported by North Pacific Research Board, and the Bering Sea Ecosystem Study (BEST) supported by the National Science Foundation addressed changes in the critical marine ecosystem of the Northern Pacific Bering Sea. More than one hundred scientists over a six year period engaged in field data collection, original research, and ecosystem modeling. The resulting 359 datasets set a new paradigm for critical information needed to answer key questions about these changes. The Earth Observing Laboratory (EOL) of the National Center for Atmospheric Research provided data management support to the BEST and BSIERP investigators, bringing their efforts together for archival into the Bering Sea Project Data Archive. This poster gives an overview of the comprehensive data management support strategy that EOL provided for the Bering Sea Project: from early involvement with the science team in order to determine their requirements; to developing a specification for metadata and documentation; from the onboard ship support and implementation of the cruise Field Catalogs on ten different ships over four years; and developing online tools for submitting metadata and data; to geolocating Local and Traditional Knowledge stories with place names on a map. The lasting legacy is a database that cross references each unique investigator dataset, providing

easy access to the metadata and data. EOL will assume long term stewardship of this unique data archive, and the poster will serve as a resource for how the archive has been conceptualized and organized.

For all details see: <http://commons.esipfed.org/node/7802>

### **Geofairy: An Open Geospatial Information Sharing Mobile Platform**

Ziheng Sun, Liping Di, Gil Heo, Eugene Yu

Center for Spatial Information Science and Systems, George Mason University, GMU, CSISS

Geofairy (<https://play.google.com/store/apps/details?id=edu.gmu.csiss.mobile.geofairy>) is a mobile app developed by the Center for Spatial Information Science and Systems in George Mason University. Geofairy succeeds to collect and display nine kinds of the mostly used geospatial information: vegetation, weather, precipitation, atmosphere, soil moisture, altitude, agriculture, land cover and land use. The information comes from different data providers, e.g. NASA, NOAA and Google. Geofairy performs as a hub gathering and reformatting information and sharing with users. It supports users to view information in three modes: table, chart and map. Through Geofairy Users could capture the situation information of interesting locations in an intuitive manner. The information from Geofairy could help make decisions in a wide range of domains like designing plans to respond agriculture drought, biodiversity crisis, climate change, disaster, disease spreading and extreme weather. Geofairy involves eight public datasets and more than 100 data layers nearly covering the whole globe, including most developing countries. For developing countries where geospatial information are lack and hard to retrieve, Geofairy allows them to monitor the measures of the environment, provide timely and accurate status reports, analyze the trends in the history and guide people in developing countries to make plans and take actions to prevent or remove the possible bad consequences.

Related Collaboration Areas: Geospatial, Information Technology and Interoperability, Products and Services, Science Software  
Tags: mobile app, geospatial information, data sharing, Decision Support

For all details see: <http://commons.esipfed.org/node/7739>

### **Proto-Examples of Data Access and Visualization Components of a Potential Cloud-Based GEOSS-AI System**

Bill Teng, Chris Lynnes

NASA GES DISC (ADNET), NASA Goddard Space Flight Center

Once a research or application problem has been identified, one logical next step is to search for available relevant data products. Thus, an early component of a potential GEOSS-AI system, in the continuum between observations and end point research, applications, and decision making, would be one that enables transparent data discovery and access by users. Such a component might be effected via the system's "data agents." Presumably, some kind of data cataloging has already been implemented, e.g., in the GEOSS Common Infrastructure (GCI). Both the agents and cataloging could also leverage existing resources external to the system. The system would have some means to accept and integrate user-contributed agents. The need or desirability for some data format internal to the system should be evaluated. Another early component would be one that facilitates browsing/visualization of the data, as well as some basic analyses. Three ongoing projects at the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) provide possible proto-examples of potential data access and visualization components of a cloud-based GEOSS-AI system. Reorganizing data archived as time-step arrays to point-time series ("data rods"), as well as leveraging the NASA Simple Subset Wizard (SSW), to significantly increase the number of data products available, at multiple NASA data centers, for production as on-the-fly (virtual) data rods. SSW's data discovery is based on OpenSearch. Both pre-generated and virtual data rods are accessible via Web services. Developing Web Feature Services to publish the metadata, and expose the locations, of pre-generated and virtual data rods in the GEOSS Portal and enable direct access of the data via Web services. SSW is also leveraged to increase the availability of both NASA and non-NASA data. Federating NASA Giovanni (Geospatial Interactive Online Visualization and Analysis Interface), for multi-sensor data exploration, that would allow each cooperating data center, currently the NASA Distributed Active Archive Centers (DAACs), to configure its own Giovanni deployment, while also allowing all the deployments to incorporate each other's data. A federated Giovanni comprises Giovanni Virtual Machines, which can be run on local servers or in the cloud.

Related Collaboration Areas: Cloud Computing, Discovery, Information Technology and Interoperability, Visualization  
Tags: GEOSS, AI, agents, data, Discovery, ACCESS, Visualization, data rods, federated Giovanni, Cloud

For all details see: <http://commons.esipfed.org/node/7822>

## **Going Beyond the Earthdata Website: Connecting with End Users Using Social Media and Webinars**

Minnie Wong, Jennifer Brennan, Jeanne Behnke, Ross Bagwell, Kevin Murphy  
NASA ESDIS, NASA Goddard Space Flight Center, CTS, NASA

In an effort to foster a two-way conversation and reach out to our diverse end-user communities world-wide, the Earth Observing System Data and Information System (EOSDIS) project has launched several new NASA Earthdata social media accounts and since May of 2013, has been hosting monthly data discovery and data access webinars. This poster will provide information about our social media efforts highlighting some key findings/metrics surrounding these efforts as well as to convey the importance and success of our monthly NASA Earthdata webinar series. Tying into the revised Earthdata content curation schema, we will be leveraging richer content and input from the Distributed Active Archive Centers, Science Computing Facilities and related Earthdata components such as Worldview, the Earthdata Code Collaborative (ECC) and the Global Imagery Browse Services (GIBS). Future web features may include user stories from the Earth science research communities, informatics articles, data discovery and data access tips, tutorials, and other data recipe "how-to's".

Related Collaboration Areas: Discovery, Education  
For all details see: <http://commons.esipfed.org/node/7847>

## **A Practical Conceptual Design of Cyberinfrastructure for Earth Sciences (D)**

Chaowei Yang, Min Sun, Erin Robinson, Zhenlong Li, Manzhu Yu

The poster reports the findings and documents developed from our NSF EarthCube project, which takes an agile/cyclic process to develop a conceptual architecture for EarthCube. We first analyzed the community needs by sorting through ~30 workshop reports to filter out the capability needs and categorize them into three parts as resource capability, enabling capability, and end user capabilities. Each of them includes a number of popular capability modules and are specified in the poster. After this analyses, we refer to 5 popular enterprise architectures to come up with a comprehensive enterprise architecture includes four volumes: volume 1 is an introduction of the overall design and includes the background, process, how to read and use the entire document sets. Volume 2 includes the details of a conceptual architecture design with use cases added. Most architecture related elements are elaborated in this volume and specific users can refer to the parts of their interest. Volume 3 is the dictionary and vocabulary structure. Volume 4 is an example about how to use the conceptual design to come up with a specific architecture for polar CI portal.

Related Collaboration Areas: Discovery, Earth Science Collaboratory, Geospatial, Information Technology and Interoperability  
Tags: EarthCube, Architecture, interoperability, layered design  
For all details see: <http://commons.esipfed.org/node/7777>

## **VegScape: an update on the large-scale crop condition and progress monitoring system**

Eugene Yu, Liping Di, Zhengwei Yang, Rick Mueller, Weiguo Han, Bei Zhang

VegScape (<http://nassgeodata.gmu.edu/VegScape/>) has been collaboratively developed as a national crop condition monitoring system by the Center for Spatial Information Science and Systems of George Mason University and the National Agricultural Statistics Service of USDA. The system is an open, Web-based, systems consists of geospatial Web services and components, including Web services following OGC specifications - Web Coverage Service (WCS) for data/product servicing, Web Map Service (WMS) for data presentation, Web Feature Service (WFS) for geographic features, and Web Processing Services (WPS) for data processing and analysis. The system enables the live link from Earth Observations (EO) to final crop condition indices and crop progress stages. The latest focused development was on crop progress modeling using time series analysis of crop condition profiles. Progressive double sigmoid models along with series of smoothing and noise-filtering have been applied in improving the crop growth stage estimation. The system currently estimates the 10 major crops in the States. VegScape reaches the stage of operational adoption in crop condition monitoring and crop progress stage estimation using primarily very high temporal resolution EO data.

Related Collaboration Areas: Geospatial, Information Technology and Interoperability, Products and Services, Visualization  
Tags: crop condition, crop progress, remote sensing, geospatial Web service  
For all details see: <http://commons.esipfed.org/node/7733>

## **CINERGI: Community Inventory of EarthCube Resources for Geoscience Interoperability**

Ilya Zaslavsky, Luis Bermudez, Raquel Calderon, Chris Condit, Jeffrey Grethe, Amarnath Gupta, Leslie Hsu, Kerstin Lehnert, Kai Lin, Tanu Malik, Burak Ozyurt, Stephen Richard, David Valentine, Thomas Whitenack  
San Diego Super Computing Center, OGC, University of California at San Diego, Columbia University, IEDA, University of Chicago, Arizona Geologic Survey

The goal of the EarthCube CINERGI (Community Inventory of EarthCube Resources for Geoscience Interoperability) project is to create a methodology and assemble a large inventory of high-quality information resources with standard metadata descriptions and traceable provenance, across geoscience domains. The inventory is compiled from metadata catalogs maintained by governmental and academic data facilities, as well as from user contributions. Once harvested into CINERGI, metadata records are processed according to harvest adapter definitions, loaded into a staging database implemented in MongoDB, and validated for compliance with ISO 19115/19139 metadata model and schema. Several types of metadata defects detected by the validation engine are either automatically corrected with help of several information extractors or flagged for manual curation. The metadata harvesting, validation and processing components generate provenance statements using W3C PROV notation, which are stored in a Neo4J database. All these components are organized into CINERGI metadata curation pipeline. The core component of the pipeline is a set of "metadata enhancers," which represent services responsible for correcting or enhancing metadata content: adding spatial extent information; adding keywords based on a collection of registered vocabularies and SciGraph annotation API, validating and correcting organization information, etc. The curated metadata records, along with provenance information, are re-published and accessed programmatically and via a CINERGI online application. The latter represents a custom search application built over Geoportal, SOLR and Neo4J. The project's website is <http://workspace.earthcube.org/cinergi>. Support from the US National Science Foundation under award NSF ICER-1343816 is gratefully acknowledged.

Related Collaboration Areas: Documentation, Information Technology and Interoperability, Semantic Web

Tags: catalogs, EarthCube, data discovery, metadata, curation

For all details see: <http://commons.esipfed.org/node/7840>