Sessions Proposals/Ideas for ESIP Data Stewardship Committee Winter Meeting 2017

(Sessions that are tagged with "Data Management Training", "Data Preservation", "Information Quality", and/or "Preservation and Stewardship")

Last Updated: 20161014

1. Title (link):Information Quality Cluster - Fostering Collaborations - http://commons.esipfed.org/node/9508
    a. Session Lead: H. K. "Rama" Ramapriyan
    b. Abstract/Agenda:
    The goals of the Information Quality Cluster (IQC) as expressed in the Strategic Plan are: 1. Bring together people from various disciplines to assess aspects of quality of Earth science data; 2. Establish and publish baseline of standards and best practices for data quality for adoption by inter-agency and international data providers; and 3. Build a framework for consistent capture, harmonization, and presentation of data quality for the purposes of climate change studies, Earth science and applications. Moving towards these goals, the IQC has been collecting and evaluating use cases to help identify issues and analyze them to arrive at recommendations for improvements in capturing, describing, enabling discovery and facilitating use of data quality information. The purpose of this session will be to collaborate with other ESIP clusters whose primary emphasis is on utilization of Earth science data for research and applications. The purpose of the collaboration will be to answer questions such as: What type of quality information is needed for their applications? Are they easy to find? Are they complete and easy to understand? What level of data quality is important for their applications (what can they "get away with"? What improvements should be made in conveying quality information?

    A brief introduction will be provided to familiarize new attendees with the IQC. Brief presentations will be made by invited panelists from other ESIP clusters such as the Disaster Lifecycle Cluster and Agriculture and Climate Cluster representing data users' perspectives, and a panelist from the data provider community. The presentation will be followed by a discussion period to identify gaps and approaches to filling the gaps and addressing issues.
    Agenda:

- Introduction to Information Quality Cluster and Status - Ramapriyan - 8 minutes
- Panelist Presentations (Total of 8 minutes)
- Agriculture and Climate Cluster - Speaker TBD - 8 minutes
- Disaster Lifecycle Cluster - Speaker TBD - 8 mintues
- CEOS/WGCV Land Product Validation - Pierre Guillevic (U of MD) - 8 minutes

- Obs4MIPS - Robert Ferraro (JPL) - 8 minutes
- Discussion - 50 minutes

2. Title: Towards systematically curating and integrating data product descriptive information for data users

http://meetings.esipfed.org/node/9514

Session Leads: Ge Peng, Nancy Ritchey, and Sean Gordon

Abstract: Complete, consistent, and easy to understand information about data products is critical for meeting data discoverability, improved accessibility and usability, and interoperability requirements.

In the BigData and Open Data Era, with ever increasing variety and number of data products, it becomes increasingly impractical to do so in a manual fashion. The most effective way to ensure the completeness and quality of metadata and description documents of data products is to curate them in a systematic, consistent, and automatic fashion based on standards, community best practices, and defined frameworks.

Efforts to meeting this goal have been carried out in various disciplines and projects. This session invites presentations to describe and share their work/progress with the ESIP community on systems, tools, frameworks, workflows, etc. that enable repositories/data centers to systematically generate and provide descriptive information about the data products to data users for improved discoverability, transparency, usability, and interoperability. Additionally this session will discuss gaps that still need to be addressed.

Draft Agenda

- NOAA OneStop Metadata Team - TBD
- ESIP Documentation Cluster - TBD
- Obs4MIPS - Ferraro, Robert
- Environmental Data Initiative (EDI) - O'Brien, Margaret
- DataONE MetaDIG team - Mecum, Bryce
- Open Floor Discussion - 30 mins

3. New platforms and techniques for strengthening ties between observations and user communities

Abstract/Agenda:
Over the past couple of years research data stewardship platforms and tools have grown substantially. Yet when we say "grown" it may not be quite in the way one might think. Instead of growing larger as a unit, a measurement of growth here is in a systems ability to connect and the perception of relevance and utility of that system to the user communities. In this session we propose to introduce several platforms and initiatives within the research data management realm that exhibit this enhanced capacity to connect with the user on their terms. These platforms are unique in their missions yet they also connect with one another in ways that advance the overall utility for the user. We would like to get a larger picture of the

current work of several of these entities, their common threads, and the specific projects with which they are engaged.
The session will include four to five individuals representing research data platform products that exemplify these qualities. These include:

- Fedora - a flexible, modular, open source repository platform with native linked data support
- Data Conservancy - building infrastructure for data curation including discovery, sharing, and enhanced preservation
- R-Map - a project to develop mechanisms to preserve the many-to-many complex relationships among scholarly publications and their underlying data, thereby supporting the continual development of scholarly communication and digital publishing
- OSF - A scholarly commons to connect the entire research cycle
- SHARE - A higher education initiative to build a free, open, data set about research and scholarly activities across their life cycle.

By leveraging the ways these entities work together in the user's favor, this in turn helps to solve the issue of too many disperate and unsustainable singular operations. As Simon Porter (2016) notes, system boundaries are disappearing and in its place emerges a "new research information citizenship" - one based on a network of contributors not confined to one specific system. Looking at these groups we can begin to see this emerging. The platforms and their utilities are of interest to ESIP members on several different levels. They represent some of the latest work in user-centered research data platforms and they also focus on openness, transparency, and ways to foster connections among different stakeholders and data user groups. In turn ESIP provides a unique venue for these platforms to explore new ways to work with the communities most engaged in working with multiple types of earth science data.

Science, Digital; Porter, Simon (2016): Digital Science White Paper: A New ‘Research Data Mechanics’. figshare.https://dx.doi.org/10.6084/m9.figshare.3514859.v1 Retrieved: 16 56, Oct 27, 2016 (GMT)