Session: IN011. Data Stewardship: in Theory and in Practice

Session description:

https://fallmeeting.agu.org/2013/scientific-program/session-search/sessions/in011-data-stewardship-in-theory-and-in-practice-2/

ESIP's "Provenance and Context Content Standard" Use cases: developing examples and models for data stewardship

Sarah Ramdeen, University of North Carolina-Chapel Hill, School of Information and Library Science
Denise Hills, Geological Survey of Alabama
Research team: Data Preservation and Stewardship Committee for the Federation of Earth Science Information Partners
Others/specific individuals???

Abstract:
Earth science data collections range from individual researchers' private collections to large-scale data warehouses, from computer-generated data to field or lab based observations. These collections require stewardship. Fundamentally, stewardship ensures long term preservation and the provision of access to the user community. In particular, stewardship includes capturing appropriate metadata and documentation--and thus the context of the data's creation and any changes they underwent over time--to enable data reuse.

But scientists and science data managers must translate these ideas into practice. How does one balance the needs of current and (projected) future stakeholders?

In 2011, the Data Preservation and Stewardship Committee (DPSC) of the Federation of Earth Science Information Partners (ESIP) began developing the Provenance and Context Content Standard (PCCS). As an emerging standard, PCCS provides a framework for *what* must be captured or preserved as opposed to describing only *how* it should be done. Originally based on the experiences of NASA and NOAA researchers within ESIP, the standard currently provides data managers with content items aligned to eight key categories. While the categories and content items are based on data life cycles of remote sensing missions, they can be generalized to cover a broader set of activities, for example, preservation of physical objects. These categories will include the information needed to

ensure the long-term understandability and usability of Earth science data products.

In addition to the PCCS, the DPSC is developing a series of use cases based on the perspectives of the data archiver, data user, and the data consumer that will connect theory and practice. These cases will act as specifications for developing PCCS-based systems. They will also provide for examination of the categories and content items covered in the PCCS to determine if any additions are needed to cover the various use cases, and also provide rationale and indicate priorities for preservation. Though the use cases currently focus on two areas, *creating* a data set and *using* a data set, the use cases will eventually cover the full data lifecycle. Currently developing a template to be used in future use case creation, the DPSC is also preparing and testing more use case scenarios.

This presentation will introduce the ESIP use cases based on the PCCS. It will at once expand stakeholder participation and show the application of these materials beyond the ESIP community in which they were developed. More information about the ESIP use case activities can be found on the DPSC wiki - http://wiki.esipfed.org/index.php/Preservation_Use_Case_Activity.