Geophysical Research Abstracts
Vol. 14, EGU2012-9594, 2012
EGU General Assembly 2012

Report on GeoData 2011 Workshop - Data Life Cycle, Integration and Citation

R. Signell (1) and P. Fox (2)

(1) US Geological Survey, Woods Hole, MA USA (rsignell@usgs.gov), (2) Rensselaer Polytechnic Institute, Troy, NY USA (pfox@cs.rpi.edu)

The U.S. GeoData 2011 was inspired by a joint NSF-USGS identification of the need to hear from the broader 'geo' community on a variety of data related matters. While increasing attention needed to be paid to full life cycle of data, in the process of preparing and scoping the workshop two other hot issues were identified: integration and citation, giving the workshop three subject areas to delve into as well as to explore connections among them. Invited participants were drawn from all 'Geo' disciplines, and beyond, from information, computer and library science, from academia, agency and commercial organizations, and from student to senior faculty/ administrators. The workshop diversity provided a rich exchange of ideas, experiences and challenges for GeoData. Many key findings and recommendations have been extracted from the detail breakout discussions and syntheses during and after the workshop. Topical categories included: metadata, standards, standards-based tools, culture, collaboration and workforce.

Key points that cut across all three-subject areas were:

- A shift is needed within agencies to provide longer-term funding support, for communities to come together, remain coherent and to enable data stewardship, integration and citation within their communities and across to other communities (to the extent possible).
- Agencies like USGS, NASA and NOAA must also play a key role in sustaining geoscience cyberinfrastructure by moving research advances into operations.
- Community-wide standards and practices should build from demonstrated successes, be widely disseminated, and tools need to be developed to support them.
- Education is critical to broader adoption. Marketing studies need to be conducted to provide the business case for full stewardship, integration and citation, and incentives are needed to encourage everyone to participate in making data integratable, citable, etc.

While technology gaps are still evident across the three topic areas, there is recognition that human factors dominate, and often limit progress and effectiveness. Organizational and resource factors that would lead to a solid understanding of the business case behind, for example, making data preservable and integratable, come at a cost that is not well documented or supported when resource (funds and people) decisions are made. This presentation will report on the data life cycle findings and recommendations from the workshop.

Recommendations:

- Create an interagency working group to foster data life cycle management practices, support coordination and informatics science initiatives, and to develop a high level shared vision and strategy for data life cycle management
- Create and fund a coordination office that works with the "working" level of agencies and academic institutions to facilitate working groups and workshops, adoption of standards and tools, and creation of sustainable archives
- Support and actively participate in existing coordination groups, and create where needed, new communities of practice in data life cycle management across agencies, academic institutions, and the private sector
- Establish an NSF working group to address the issue of incentives and cultural change needed to facilitate implementation of data life-cycle management

http://tw.rpi.edu/web/Workshop/Community/GeoData2011, #geodata2011