AGU FALL MEETING
San Francisco | 14 – 18 December 2015

Abstract ID: 77357
Final Paper Number: PA53A-2231
Session Date: Friday, 18 December 2015; 13:40 - 18:00
Session Number and Title: PA53A: An Interoperability Challenge for the Geosciences II Posters

Academic Research Library as Broker in Addressing Interoperability Challenges for the Geosciences

UNM COLLEGE of UNIVERSITY LIBRARIES & LEARNING SCIENCES

Plato L. Smith II, CLIR/DLF Postdoc Fellow
University of New Mexico
psmith4@unm.edu

ABSTRACT

The National Science Foundation (NSF) now requires a data management plan (DMP) for researchers seeking funding. Even if a funding agency does not require a data management plan, developing and implementing quality research data management and curation practices/workflows are needed across all disciplines. Academic Research Libraries (ARL) Data Services such as those institutions that participated in the ARL SPEC Kit 334: Research Data Management Services (July 2013) http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/ can help researchers and scientists with research data collection, management, and curation.

Academic Research Libraries have expertise in providing long-term discovery and access through the use of metadata standards and provision of access to research data, datasets, and publications via institutional repositories. Metadata crosswalks, open archival information systems (OAIS) [Fig. 1.a] [4], trusted data repositories [7], persistent URL, data sharing and publications, interoperability, and data management and curation compliant with evolving funding agencies' requirements [5] [9] are common digital library discipline components. These digital library discipline components complimented by the USGS Community for Data Integration (CDI) Science Support Framework (SSF) along with interactions with Open Geospatial Consortium (OGC) Programs [6] and success metrics for virtual research environments (VRE) [2] research development can contribute to advancing data management and curation for the Geosciences.

METHODOLOGY

i. (1) Attended 2015 DLF Forum and 2015 DLF e-Research Network Webinars; (2) Conducted an EarthCube Funded Projects survey (March/April 2015); (3) Attended DataONE, ESIP, USGS CDI working groups Webinars and the EarthCube (EC) RCN EC3 2015 Field Trip for Geoscientists; (5) Participated in EarthCube Gap Analysis Working Group and Coauthored a journal article.
ii. Searched re3data – **Subject**: **Geosciences (including Geography)**; **Filter**: Certificates, Open Access, Persistent Identifier (DOI); **Content Type**: Archived data, Plain text, Scientific and statistical data format, Software applications, Standard office documents, and structured text; **Country**: United States; Result = **USGS Science Data Catalog [Fig. 1.d] [7]**
iii. Participant observer, unobtrusive observations, working group and face-to-face discussions, and survey of researchers from item i. yield the following challenges:
- Lack of Metadata (Data Producer and Consumer), Best Practices, and utilization of Standards (Published, Community)
- Features/Functionality, Interoperability, Data Collection & Integration, Semantics issues
- Desktop/Mobile Applications, Web Services (e.g. APIs, W3C, SOAP, RESTful, etc.)
- End-to-End development (e.g. prototype/end of project funding)

Communications, Collaborations, & Compliance

Criterion in OSTP Public Access Memo 2013		NSF's Public Access Plan 2015	Research Data Services (RDMS)	Figures 1.a - 1.f
1	Policy Principles, including consultation with stakeholders	1.0; 1.1; 1.2; 7.0; 9.0; 10.0	Campus/Library IT, HPC, IR, OVPR consultations and collaborations	1.a, 1.b, 1.c, 1.d, 1.e, 1.f
2	Agency Public Access Plan	2.0; 3.0; 4.0; 5.0; 6.0; 7.0	Data and Open Access Policies; RCR/DMP Outreach and Training	1.a, 1.b, 1.c, 1.d, 1.e, 1.f
2.a	Strategy for leveraging existing archives and fostering public/ private partnerships with scientific journals	7.1.1; 7.1.2; 7.2.1; 7.3.1; 7.4.1; 10.0	Standards-based repositories (e.g. OAIS, DSA, re3data, TDR); Consortia, Membership, and Research networks partnerships	1.a, 1.b, 1.c, 1.d, 1.e, 1.f
2.b	Strategy for improving the public's ability to locate and access digital data	7.3.1; 7.4.2	OA IR and data repositories; Data citation, linking, and publication best practices/guidelines (e.g. DOI, Force11, Github, OGC, ORCiD, Zenodo)	1.a, 1.b, 1.c, 1.d, 1.e, 1.f
2.c	Approach for optimizing search, archival, and dissemination features that encourage innovation in accessibility and interoperability	7.4.1	Develop, leverage, and promote web, machine-readable data, Non-proprietary format, RDF standards, and Linked RDF data & information exchange across domains - Linked Open Data (LOD)	1.a, 1.b, 1.c, 1.d, 1.e, 1.f

Table 1: Sample of 2013 OSTP Memo crosswalk to 2015 NSF Public Access Plan mapped to Research Data Services (RDMS) and Figures 1.a – 1.f

CONCLUSIONS

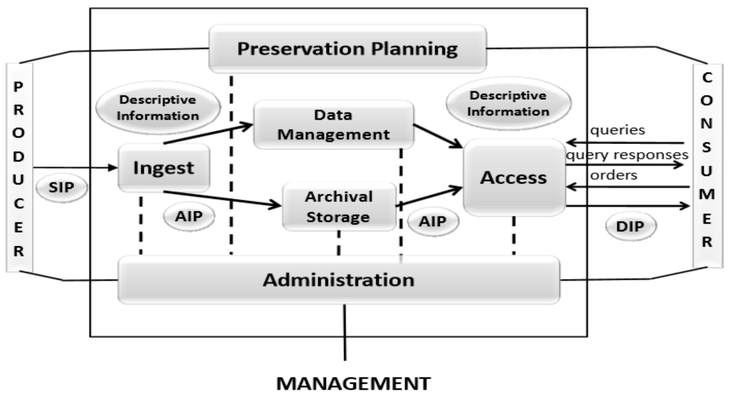
Adopting standards such as Open Archival Information System (OAIS) [Fig 1.a] principles into the data management and curation lifecycle processes allows the capture of different types of metadata as a standard part of the data preservation, management, and sharing workflow for data repositories [3]. Developing standards-based workflows facilitate metadata crosswalks and repository interoperability that enhances data access, discovery, and knowledge transfer. Funded projects and programs that incorporate general [Fig 1.a] and discipline-specific [Fig 1.c] data management curation models and/or interact with Open Geospatial Consortium (OGC) interoperability programs [Fig 1.b] [6] in the development of projects and prototypes while also utilizing existing best practices and standards in the development of a research data repository [Fig 1.d] develop foundational, syntactic, and semantic interoperable workflows [Fig 1.e] capable of enabling successful and useful metrics for evaluation [Fig 1.f] that benefits stakeholders, researchers, and users [2] that is continually needed across all domains.

REFERENCES

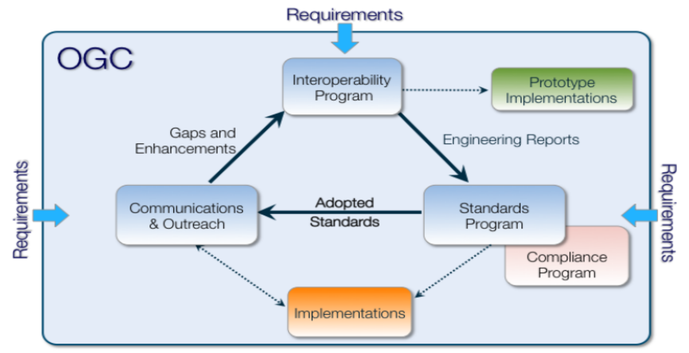
1. Brandt, S. D. (2015). DLF E-Research Network 2015 Webinar on "Introduction/Research Data Management Services in Academic Libraries" for 2015 DLF E-Research Network Cohort, May 13, 2015.
2. Buddenbohm, S., Enke, H., Hofmann, M., Klar, J., Neuroth, H., & Schwiegelshohn, U. (2015). Success Criteria for the Development and Sustainable Operation of Virtual Research Environments (VRE). D-Lib Magazine, September/October 2015, v.21 (*9/10*). http://www.dlib.org/dlib/september15/buddenbohm/09buddenbohm.html.
3. Mookerjee, M., Vieira, D., Chan, M. A., Gil, Y., Pavlis, T. L., Spear, F. S., & Tikoff, B. (2015). Field Data Management: Integrating Cyberscience and Geoscience. Earth & Space Science News (EOS), 96(20), pp. 18-21.
4. OAIS. (2012). CCSDS Reference Model for an Open Archival Information System (OAIS). Recommended Practice, CCSDS 650.0-M-2, Magenta Book, June 2012. http://public.ccsds.org/publications/archive/650x0m2.pdf.
5. NSF. (2015). NSF Public Access Plan: Today's Data, Tomorrow's Discoveries (NSF 15-52). http://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf.
6. Percivall, G., Idol, T., Alameh, N. & Harrison, J. (2015). Innovation in OGC: The Interoperability Program. ISPRS Int. J. Geo-Inf. 2015, 4(*4*), 2362-2378; doi:10.3390/ijgi4042362. Figure 9, pg. 2374.
7. re3data. (2015). http://www.re3data.org/. (Methodology ii search result is http://tinyurl.com/q9jmuha)
8. USGS Community for Data Integration Science Support Framework. (2015). http://www.usgs.gov/cdi/about.html.
9. Van de Sompel, H. (2015). Reminiscing About 15 Years of Interoperability Efforts. D-Lib Magazine, November/December 2015, v.21(*11/12*). http://www.dlib.org/dlib/november15/vandesompel/11vandesompel.html.
10. White House OSTP. (2013). Memorandum on Increasing Access to the Results of Federally Funded Scientific Research. http://tinyurl.com/qhzuuy7.

Using Preservation, Interoperability, and Science Support Framework Models for Enhancing Data Curation, Management, Sharing, and Metrics Practices

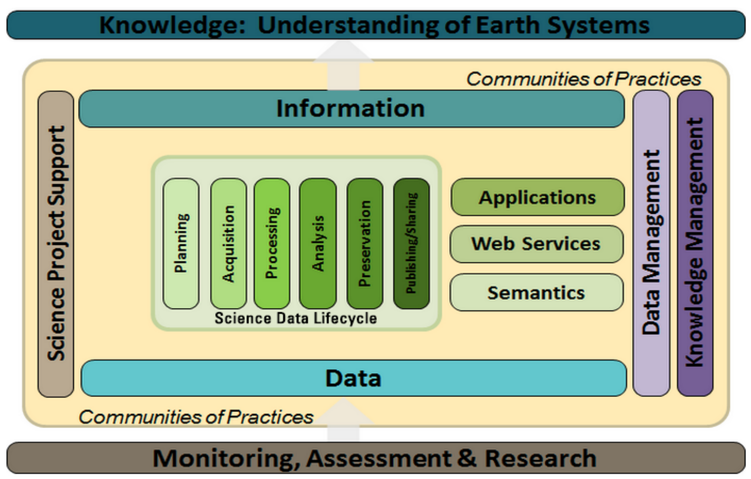
OAIS Reference Model (1.a) [4]



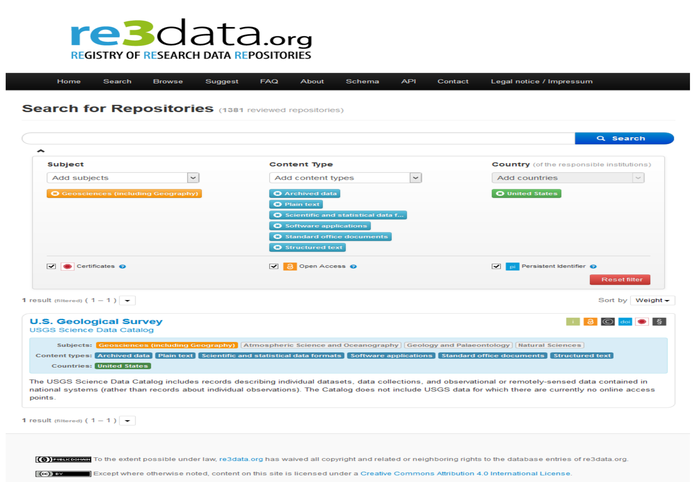
Interactions with OGC Programs (1.b) [6]



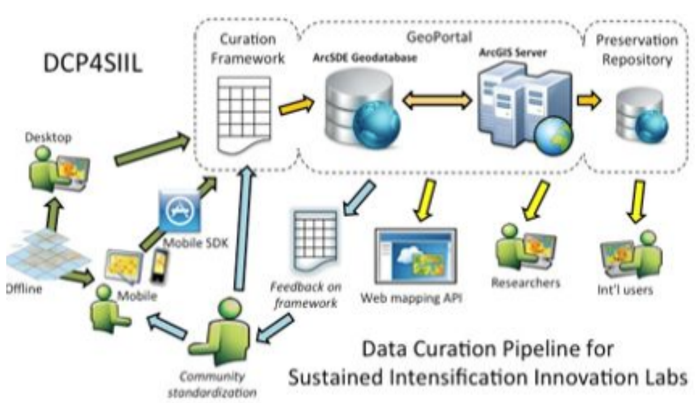
USGS CDI Science Support Framework (1.c) [8]



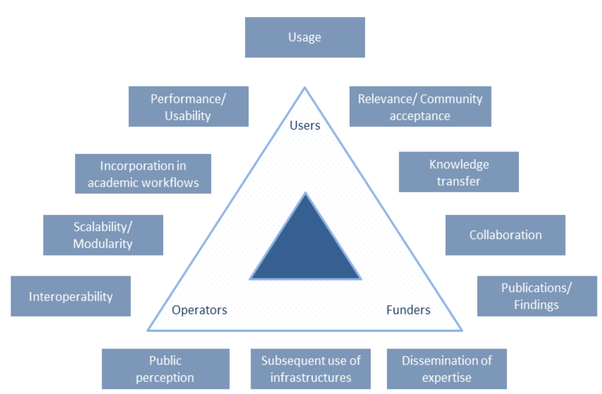
Trusted Data Repository (1.d) [7]



Research Data Management Workflows (1.e) [1]



Success Metrics (1.f) [2]



Figures 1.a – 1.f: OAIS Reference Model [1.a], Interactions with OGC Programs [1.b], USGS CDI SSF [1.c], trusted data repository for Geosciences (USGS) [1.d], Data Curation Pipeline for Labs [1.e], and VRE Success Metrics [1.f]