The Community for Data Integration (CDI): connection and collaboration with the Research Data Alliance

Leslie Hsu, U.S. Geological Survey, Denver, CO, United States (lhsu@usgs.gov)

How can two data organizations use each other's networks and outputs?

The Community for Data Integration (CDI) at the U.S. Geological Survey is a dynamic community of practice of over 700 members. The CDI learns together to develop capabilities that increase the value of scientific data and information. Although more focused in disciplinary scope to geological, hydrological, and biological data, the CDI has many similarities to the Research Data Alliance (RDA), including overlapping and intersecting topics of interest.

This poster highlights potential and existing connections between CDI and RDA after an analysis of all RDA groups and outputs. Potential connections include presentations to share knowledge on common interests and challenges, leveraging outputs and adopting recommendations, collaboration on future projects, and increased dual membership.

Community for Data Integration (CDI) Groups	Related RDA Working & Interest Groups	Related RDA Recommendations & Outputs
All CDI A dynamic community of practice learning together to develop capabilities that increase the value of scientific data and information. Meets virtually on a monthly basis to hear about new opportunities, tools, and projects.	• Data Rescue IG • Physical Samples and Collections in the Research Data Ecosystem IG • Repository Platforms for Research Data IG • Reproducibility IG • Research Data Provenance • Small Unmanned Aircraft Systems' Data IG • Virtual Research Environment IG (VRE-IG)	• 23 Things: Libraries For Research Data • Framework for Summer Schools in Data Science and Cloud Computing
CDI Data Management Working Group A group to develop mechanisms for incorporating data management into USGS science and educate scientists of its value. Includes the Metadata Reviewers Community of Practice.	• Active Data Management Plans IG • Data Discovery Paradigms IG • Data Foundations and Terminology IG • Metadata IG • RDA/WDS Certification of Digital Repositories IG • RDA/WDS Publishing Data IG	• An open, universal literature-data cross-linking service Bibliometric Indicators for Data Publishing • BioSharing Recommendations: data repositories, standards and policies in the life, biomedical and environmental sciences • Matrix of use cases and functional requirements for research data repository platforms • Metadata Standards Directory • Persistent Identifier Type Registry • Scalable Dynamic-data Citation Methodology • Workflows for Research Data Publishing: Models and Key Components
CDI Semantic Web Working Group A group exploring semantic web technologies as possible methods for USGS data and information projects.	• Vocabulary Services Interest Group	
CDI Earth Science Themes Working Group A group to share methods, data, software, and conceptual models within theme-specific applications.	• Global Water Information IG	
CDI Tech Stack Working Group A group learning about the "stack" of technology that can be used to promote the management and use of data.	• Software Source Code IG	
CDI Software Development Cluster A community for USGS software developers and other interested parties to discuss software release protocols and policies, development best practices, software metadata, and software libraries, packages, and tools.	• Software Source Code IG	

Summary and Next Steps

CDI as a whole and the CDI Data Management Working group have the most potential connections with current RDA groups and outputs, although there are strong connections between other CDI-RDA pairs. Some beneficial dual-membership between the organizations already exists. We focused on near-term connections and use of outputs, so some of the newer RDA groups are not listed here.

The CDI is in the process of reaching out to points of contact in RDA to see if they would like to present their recommendations or other findings to the CDI, or if they have interest in getting in touch with specific USGS groups or data asset representatives.

If you are part of an RDA group that you think has potential connections with one of the CDI groups listed here, get in touch with us at lhsu@usgs.gov.

Learn more about the Community for Data Integration at bit.ly/usgs-cdi-wiki.

Research Data Alliance Plenary 10, Montreal, September 2017