



Federation of Earth Science
Information Partners

MAKING DATA MATTER

RDA and ESIP:

How can we work together?

- Ruth Duerr

July 12, 2013



National Snow and Ice Data Center
Supporting Cryospheric Research Since 1976



Working Groups and ESIP

RDA Working Groups

- Data Type Registries
- PID Information Types
- Standardization of data categories and codes, working specifically with the ISO 639 (human languages)
- Data Foundation and Terminology (pending)
- Practical Policy (pending)
- Community Capability Model (in review)
- Data Citation: Making Data Citable (in review)
- Metadata Standards (in review)

ESIP Collaborations

- Air Quality
- CF Standards
- Climate and Agriculture
- Climate Education
- Cloud Computing
- Preservation and Stewardship
- Data Management Training
- Decisions
- Discovery
- Drupal
- Earth Science Collaboratory
- Energy and Climate
- Education
- Geospatial
- Information Quality
- Information Technology and Interoperability
- Internal Education
- Products and Services
- Semantic Web
- Visualization



Working Groups and ESIP

RDA Working Groups

- Data Type Registries
- **PID Information Types**
- **Data Foundation and Terminology (pending)**
- Practical Policy (pending)
- Community Capability Model (in review)
- **Data Citation: Making Data Citable (in review)**
- Metadata Standards (in review)

ESIP Collaborations

- CF Standards
- Cloud Computing
- **Preservation and Stewardship**
- Discovery
- **Earth Science Collaboratory**
- Information Quality
- **Information Technology and Interoperability**
- Semantic Web

Data Types Registries

The Data Type Registries Working Group will compile a **set of use cases** for data type use and management. It will **identify** and distinguish among **existing 'type registry' efforts** and their potential interaction with this group and **formulate a data model and expression for types**. It will also **design a functional specification for type registries** and propose a **federation strategy among multiple type registries** at both technical and organizational levels.



PID Information Types

This Working Group focuses on cross-community concerns regarding **Persistent Identifiers** which are the core of proper data management and access..

- Scope the type standardization effort and its boundaries
- Define core PID information types
- Define the role of structural elements, such as collections, in combination with the information types
- Define a profiles mechanism and demonstrate its scope with some initial exemplar profiles
- Provide a first API prototype that demonstrates a possible implementation of these elements and accompanying rudimentary services



Data Foundation and Terminology

The Data Foundation and Terminology Working Group describes a basic **abstract data organization model** which can be used to derive a **reference data terminology** that can be used across communities and stakeholders to better synchronize conceptualization, enable better understanding within and between communities and stimulate tool building



Practical Policy

The Practical Policy Working Group will assemble a **collection of production and research policies**, analyze the submitted policies to **identify best practices**, and **promote** the formation of **policy-based data management systems**. **Categories of computer actionable policies will be defined** to simplify policy-based interoperability, with the goal of identifying a minimal desired set for policies for each category that can be used as a starter kit. Example categories include collaboration, publication, preservation, curation, and data processing.



Community Capability Model

The purpose of the RDA CCM Working Group (WG) is to **collect, validate and publish a range of data-centric “capability profiles”** to enhance inter- and intra-domain interoperability and catalyse RDA data-sharing goals. Various barriers (technical, social, legal, ethical), currently make this difficult to achieve beyond established domain community boundaries, limiting the penetration and impact of **data-intensive research**. The CCM-WG also has potential to adopt a foundational role in synthesising outputs from other RDA WGs.



Data Citation: Making Data Citable

The RDA Working Group on Data Citation (WG-DC) aims to bring together a group of experts to discuss the **issues, requirements, advantages and shortcomings of existing approaches** for efficiently **citing subsets of data**. The WG-DC focuses on a narrow field where we can contribute significantly and provide prototypes and reference implementations.



Now - how can we work together?

Suggestions:

- Liaisons between groups
- Other suggestions?





Interest Groups and ESIP

RDA Interest Groups

- Agricultural Data Interoperability
- Brokering
- Legal Interoperability
- Metadata Standards Directory
- Preservation e-Infrastructure
- Repository Audit and Certification
- Structural Biology
- The Engagement Group
- Big Data Analytics
- Data in Context
- Defining Urban Data Exchange for Science
- Marine Data Harmonization
- Publishing Data
- UPC Code for Data
- Digital Practices in History and Ethnography

ESIP Collaborations

- Air Quality
- CF Standards
- Climate and Agriculture
- Climate Education
- Cloud Computing
- Preservation and Stewardship
- Data Management Training
- Decisions
- Discovery
- Drupal
- Earth Science Collaboratory
- Energy and Climate
- Education
- Geospatial
- Information Quality
- Information Technology and Interoperability
- Internal Education
- Products and Services
- Semantic Web
- Visualization



Interest Groups and ESIP

RDA Interest Groups

- **Agricultural Data Interoperability**
- **Brokering**
- Legal Interoperability
- **Metadata Standards Directory**
- **Preservation e-Infrastructure**
- **Repository Audit and Certification**
- The Engagement Group
- Big Data Analytics
- **Data in Context**
- Defining Urban Data Exchange for Science
- Marine Data Harmonization
- **Publishing Data**
- UPC Code for Data

ESIP Collaborations

- Air Quality
- CF Standards
- **Climate and Agriculture**
- Cloud Computing
- **Preservation and Stewardship**
- Decisions
- **Discovery**
- Earth Science Collaboratory
- Energy and Climate
- Geospatial
- Information Quality
- Information Technology and Interoperability
- Semantic Web