RDA in Engineering

Adoption and Outputs session, Plenary 14
Relevant Groups represented

- Research Data Management in Engineering IG - Paula Martinez Lavanchy and Sibylle Hermann
- Persistent Identification of Instruments WG - Rolf Krahl
- International Materials Data Registries WG / RDA/CODATA Materials Data, Infrastructure & Interoperability IG - Laura Bartolo and Ray Plante
- Physical Samples and Collections in the Research Data Ecosystem IG - Sarah Ramdeen
Also at the table

• Ana Slavec – RDA Europe 4.0 Ambassador for Engineering / Renewable Materials
• Shaun de Witt – RDA Europe 4.0 Ambassador for High Energy Physics / Fusion
• Stefanie Kethers – RDA Secretariat
Other Relevant Groups

- Chemistry Research Data IG
- Research data needs of the Photon and Neutron Science community IG
- Data Versioning WG
- Federated Identity Management IG
- Preserving Scientific Annotation WG
- Data Citation WG
- Software Source Code identification WG
- Preservation Tools, Techniques, and Policies IG
- All Metadata IGs and WGs
International Materials Resource Registries: Outputs and Follow-on Activities

Raymond Plante -- NIST
Final Report of Outputs

- Presented at P11 (Berlin)

- Internal WG review of Final Report Document:
  - Current version is RC4
  - Please send comments to working group mailing list (imrr-wg@rd-groups.org) by 20 November 2019

- To be submitted to secretariat for RDA-wide RFC afterward.
Output Summary:

Registry Federation Framework

Data Centers

- Local Publishing Registry
  - Portal
  - Dataset
- Local Publishing Registry
  - Portal
  - Database
  - Dataset

Data Collections

- Full Searchable Registry
  - e.g. operated by NIST
  - Harvest (pull)
- Full Searchable Registry
  - e.g. operated by CHiMaD
  - Harvest (pull)
- Small Data Providers
  - Manual Entry
- Data Repository
  - Materials Data Facility
Output Summary:

Registry Federation Framework

- **Requirements**
  - Resource metadata exchange protocol
    - Identifiers
    - Distinguish between own records and those harvested from other sources
    - Communicate when resources are no longer available
    - Require minimal record validation
    - *Our implementation: OAI-PMH. (Others: Linked Data Platform (LDP), ResourceSync, ...)*
  - Common Metadata Schema/Format
    - Openly defined
    - Associate a globally unique identifier
    - Validate-able
    - Low-impact evolution mechanism (e.g. extensions)
    - *Our implementation: XML, XML Schema*

- **Architecture, Recipe for registry interaction**
Output Summary:

Metadata

- XML Schema
  - Extension mechanisms
  - Different Types of Resources:
    - Data Collections  Databases  Software  Informational Sites  Organizations
    - Different types can have different data associated with them
    - New types can be defined. (Semantic Asset)
  - Applicability to different domains
    - Place to include domain-specific metadata
    - Can support multiple domains simultaneously

- Materials Science Vocabulary
  - 3-tiered subject terms
  - Drives faceted browsing
  - SKOS definition available
Output Summary:

Working, Populated Registry Federation

- Implementation: NIST Materials Resource Registry
  - Adaptation of the NIST Configurable Metadata Curation System (CDCS)
  - https://github.com/usnistgov/MaterialsResourceRegistry

- Two instances:
  - NIST: https://materials.registry.nist.gov/
  - CHiMaD/MDF: https://mrr.materialsdatafacility.org/

- Over 350 records
Post-WG Activity:

Supporting Semantic Assets for MSE

- Adding “Semantic Asset” as a resource type to MMR
  - Cover vocabularies, ontologies, types, registries, ...
  - To encourage sharing use across continents

- MSE Vocabulary Use & Maintenance
  - Elsevier pilot: considering tagging MSE journal articles with MRR vocabulary terms
  - MatVoc pilot: MRR Vocabulary used as starting point for building a community sustainable/evolvable vocabulary

- MDII Task Groups
  - Expanding impact of vocabularies and registries
  - Task Group on Materials Ontologies
Post-WG Activity:

Enhancing Discovery

- Additional registry instances in the world
- NIST: Software improvements
  - Improving usability
  - Stronger support for PIDS and PID resolution
- Deep Discovery: leveraging data provider search tools
  - Register search services
  - Tools can pass search queries to a repository’s or database’s search service for retrieving dataset/measurement-level results
Persistent Identification of Instruments WG

Rolf Krahl

RDA P14, Helsinki, 25 October 2019
Persistent Identification of Instruments

Persistent Identification of Instruments WG

The Persistent Identification of Instruments Working Group seeks to explore a community-driven solution for globally unique identification of measuring instruments operated in the sciences.

- Create persistent identifiers (PIDs) for instruments.
- Better track the scientific output of an instrument by journal articles and datasets referencing the instrument.
- Allow automatic aggregation of information: richer metadata for datasets if information about the instrument used to collect the data can be retrieved.
- Add another relevant node to the PID Graph.
PIDINST WG Activities & Outcomes

- Collected 15 use cases from institutions interested in using PIDs for instruments.
- Formulated a schema for the metadata describing an instrument to be registered with a PID.
- Explored potential solutions with PID infrastructure providers. Identified two candidates:
  - ePIC (Handles)
  - DataCite (DOIs)
- Tried to map our metadata schema onto DataCite Schema. Found this to be feasible.
- A recommendation paper is work in progress.
- First adopters.
IG RDM in Engineering

Co-Chairs
Daniela Hausen, Paula Martinez-Lavanchy, Nikola Vasiljevic and Jonathan Peters

Aims
• Building the bridge between engineers, infrastructure, engineering associations and industry
• Address RDM-related topics in engineering science like the implementation of the FAIR principles, open research, DMPs, metadata schemes, ...
• Involve and work with different engineering subdisciplines on RDM-related topics
Starting Topics

- Engineering and Open Science
- Engineering-specific DMP
- Wind energy data annotations
- Metadata scheme EngMet and Software management

Involvement in RDA

- Most RDM topics are important on an international level.
- Learn from other disciplines and adopt it to engineering science
- Adopt generic RDM topics and make it engineering specific
- Collaborate with other IG and WGs in RDA