Working Group 1: MDCS, DSpace, MRR & MDF

Ian Foster
Ben Blaiszik
Ed Seidel
Chandler Becker
Raymond Plante
Zachary Trautt
Sharief Youssef
Laura Bartolo
Significance of WG’s Focus

• Facilitate generated materials data to be stored, registered, made discoverable, and made accessible to the materials community.

• Leverage existing efforts in the materials community so that we can work together.

• Facilitate interoperability between resources.

• Provide a framework for sub-groups in the materials community to develop customized schemas and start making use of them more broadly.
Summary of WG’s Goals

• Have a common core or minimal set of metadata tags. Ex. Dublin Core

• Allow for the common core to be extensible, adaptable, flexible enough for a domain such as materials science to use.

• Agree on granularity of metadata exchange.

• Allow for interoperability. Agree on a protocol for exchanging records.

• Provide a common authentication system.
Summary of WG’s Goals

- Develop a model for community engagement.
  - Collaboration between users
  - How to engage the current generation and the next generation?
- Leverage what other folks are already doing in the materials domain
- Facilitate control over data access (Ex. publically available or login required)
Summary of WG’s Goals

- Agree on a publication workflow.
  - Provide moderation?
  - Provide versioning of records
- Agree on policies for data management
  - How long to store the data?
- Provide for advanced search
  - Amazon-like facet search / refinement
- Make use of the RESTful APIs
- Push notification for updates
Technical Requirements/Needs

• Support for the OAI-PMH protocol
• Provide a RESTful API
• XML v.s. JSON
  – A metadata framework that can go between the two formats
• Common Metadata Schema as a Standard
  – Extensible, flexible, adaptable
• Common authentication system
Collaborations/Synergies

• Other Working Groups? Some potentials:
  – WG 2 - Experimental Data (how to describe and store?)
    • Having a common broad metadata schema
    • Build on the base metadata schema
    • Storage, Discovery, and Access
  – WG 3 - Schemas for Polymer Nanocomposite Data?
  – WG 4 - Natural Language Processing?
  – WG 5 - DFT - Density Functional Theory
    • Distributed communities and efforts. Facilitate interconnection of all the various efforts.
  – WG 6 - Building CALPHAD proto-databases?
Collaborations/Synergies

• *Materials Data Curation System*
• *Materials Resource Registry*
• *Prototype Materials Resource Description Schema*
• *Materials Data Facility*