Scope of the task group for Semantic Assets for Materials Science

Contacts: Vasily Bunakov vasily.bunakov@stfc.ac.uk
          Zachary Trautt   zachary.trautt@nist.gov

The group is established within the existing RDA Vocabulary Services Interest Group https://rd-alliance.org/groups/vocabulary-services-interest-group.html as an informal "task group" united by a common professional interest among informaticians concerned about data management for Materials Science – in particular, about semantic assets of all kinds: vocabularies, ontologies, metadata schemes, glossaries, lists of common terms.

The scope below is intended only as a "framework" to better structure discussions in the group, when in fact discussions can be about anything of a common professional interest:

- Building an inventory (up to the point of having it as a running IT service) of existing semantic assets – vocabularies, controlled terms lists, metadata schemes - for Materials Science. This can include not only vocabularies about materials per se but also cover adjacent topics, say instrumentation and chemistry. We have to be reasonable about the scope to not over-inflate it, but having identified useful pick-ups from adjacent areas will be good for Materials research and for Materials informatics.
- Monitoring technology for vocabularies building and vocabularies maintenance / updates / curation in Materials domain. This can be harvesting terms from literature (and normalizing them), or involve human experts – whatever can facilitate vocabularies creation and their proper maintenance.
- Monitoring use cases and actual practices for vocabularies (and other semantic assets) application in Materials domain. This includes using them in the actual IT services built by publicly funded projects. This also includes semantic assets "embedded" in commercial products such as Springer Materials database where access to data is protected – yet the structure of descriptions (metadata) is open, so that semantics / vocabularies they use can be monitored anyway.
- Discussing forms of representation / publishing for vocabularies (and other semantic assets). Striving for the ubiquitous Linked Data is all right but in reality, certain use cases may prefer other representations, or not be ready for the Linked Data adoption.
- Discussing interoperability between vocabularies: a possibility for crosswalks or sensible links between terms from different vocabularies. This can potentially result in significant contribution to interoperability between IT services, too.