IGSN e.V. General Assembly 2016 - Pre-meeting Technical Meeting
Sunday, December 11, 2016
Room Sierra J, Marriott Marquis, San Francisco, CA, USA
Notes from technical meeting prior to G.A.

Present: Lesley Wyborn, Irina Bastrakova, Jens Klump, Anusuriya Devaraju, Kerstin Lehnert, Anders Noren, Nancy Todd, Bob Arko, David Arctur, Steve Richard, Megan Carter

1. Anu showed a cleaned up version of the IGSN Descriptive Metadata and we discussed each element.
2. Parent Identifier and Collection Identifier elements were removed (handled by the related identifier element). Controlled vocabularies were all decoupled. Registration type is new (PhysicalSample, SampleCollection, or SamplingFeature) - we can now distinguish between them.
3. Pool Party Thesaurus (vocabulary manager) - GUI multiple accounts for multiple vocabularies, all will use the central vocabulary service (license is expensive)
4. Now two ways to specify locations (geometry + sridType or toponym).
5. Features like boreholes, outcrops, sites are assigned IGSNs
6. How should collections of IGSNs appear in a paper?
7. Copernicus now has an assets tab.
8. For collections (we might call them aggregations or groups, rather than collections), the IGSN of a specific collection (aggregation/group) cannot change. There can be a new version of the group that can get a new IGSN if samples are added to that aggregation, etc.
9. Kerstin suggested that we need to have a working group on writing up guidelines to allocating agents - these are requirements that, if you want to become an AA, you need to follow (persistent access to metadata, etc.) We need business rules.
10. Location is optional. A sample can have both a toponym and a lat/long. Should the toponym come from a specific gazetteer? Jens does not want to create an IGSN for every sampling feature.
11. WKT, GML
12. Do we want to adopt the ANDS top level vocabularies? (PhysicalSample, SampleCollection, or SamplingFeature) Bob suggests grabbing these three terms and internalize them. IGSN.org needs to support three solid URIs for those terms.
13. David had questions about sampling methods and collection types, where there are already expansive vocabularies in existence
14. Name has been renamed "Title"?
15. RelationTypes - proposal is to reuse DataCite relation types, but add a few. Bob also argues that these vocabularies should be internalized into the IGSN.org space
    a. Jens pointed out that there is duplication here because this is also included in the registration metadata. It should be in one of the two. → in other words, do those relationships belong in the birth certificate or not?
16. There are relationships between some samples (a diamond that came from a rock), but you can
17. Kerstin said that when we are identifying a group of samples in a publication, that is very different from identifying all of the samples in a specific collection at a museum. We should ask SciColl what they are doing.
    a. What was the intention of collecting those samples? If we think that way, we have to consider collections dynamic. They can change over time. So, therefore, you have to specify date as well. **Overlying use case has to be being able to reproduce the analytic results in a particular publication. But we must consider cases where**

> **there will be no publication.** Groupings are only static. All of the temporal relationships should be incorporated elsewhere. Splits of the samples in the group do not become part of the group.
>
> b. Kerstin reported that Brooks Hanson is determined to use one of the AGU Journals to enforce the use of IGSNs (G-Cubed?)

18. What about a composite sample where there are multiple parents for a single sample? (multiple soil samples pooled together to create a single new sample)
19. Combine all agents under roles? How is contributor different from collector?
    a. Anu says collector and curator are separate in CSIRO.
20. Should we make the master IGSN schema look like the CSIRO one?
21. Supplemental metadata is a URI to another record, where metadata exists elsewhere.
22. SampleAccess should be made clearer. → Change to IsMetadataPublic
23. RelatedIdentifiers → Change to RelatedResourceIdentifier
24. ResourceType is the physical form of the sample whereas registrationType is whether it is a physical sample, sample collection, or sampling feature. We would want to make sure that only some resourceTypes are applicable for some registrationTypes. Anu says we can create a grouping of enumerations (possible sample types) for each registration type (XSD 1.1 assert?)
25. Supplemental metadata can have one or more records and each record can have title and metadata URI as child elements.

The following changes were applied to the IGSN Description Schema on 14.12.2016:
1. removed unused controlled vocabs from the /include folder, e.g., collectiontype
2. locations element was updated (0-M)
3. The registrationType.xml - the enumeration now includes the igsn prefix, e.g., http://schema.igsn.org/vocab/PhysicalSample
4. Contributor type list was updated based on datacite schema v4.0
5. sampleAccess was renamed as isMetadataPublic
6. relationType - some examples from the datacite schema were included
7. supplementalMetadata>metadata now has two child elements (metadataTitle, metadataURI)

David Arctur-- questions:
Profiles for extended metadata content. SESAR-- what are sesar requirements vs what are IGSN requirements.
Is a sample still in existence and how much is left? CSIRO has this with eventType. This should happen at the allocating agent level.

Should birth certificate have the body that you are from? (Earth or the Moon or something else) How do you traverse the universe and find just the lunar samples?