

Notes from Discussion: GBIF, 21/22 March 2013

Participants/ Contributors:

Eamonn O’Tuama (GBIF)

Hannu Saarenmaa (EU-BON)

Donald Hobern (GBIF)

Oliver Meyn (GBIF)

Tim Robertson (GBIF)

Wim Hugo (SAEON)

# GBIF Services

The aim of this discussion was to determine the basis of more comprehensive use of GBIF data in other contexts. As an example, SAEON wants to establish periodic snapshots of Southern African (and possibly African) data sets with a view to creation of SensorWeb and OGC WxS services for use by its portals and communities.

Skype

Oliver Meyn described GBIF development in respect of services aimed at large download requests. This will be directly usable to support the SAEON requirement.

1. Services will be at <http://uat.gbif.org> (initially).
2. User account – determines access to services – OM will provide details.
3. Include WG 8 and AIP-6 into GBIF beta testing.
4. Contact point: OM to confirm when and let WH, HS know.
5. WH: set up the data environment in the SAEON AIP framework.
6. WH: Ask OM for download sample (zip file).
7. Follow up again start of June 2013.

# AIP/ GBIF Deliverables

The discussion aimed to determine which of the GBIF deliverables, contributed at the Asilomar discussions, would be directly usable in the GEOSS AIP-6.

1. Only some deliverables are to be included into the AIP.
   1. This has to be described in more detail in the AIP implementation (WH, HS, EoT) - end April.
   2. The basis on which GBIF participates must be clear to all stakeholders, WH and EoT to confirm this in a short document – end April:
      1. GBIF provides services and e-mail support.
      2. Third parties (WG 5 members) will test, implement, take responsibility.
      3. EoT is the contact point.
2. GBIF is a fundamental contributor of data services to AIP (either directly or after repackaging by others – SAEON, etc.)
3. WH to confirm the interaction with the GCI/ GEOSS Broker – Stefano Nativi (SN) – hopefully an API. Start of April.
4. Confirm milestones with SN, publish detailed AIP work plan end of April. (WH)

# Extending DwC

1. There are three processes aiming to extend DwC that are interlinked:
   1. ‘BioData’:
      1. NIWA (Jochen Schmidt (JS) initiated discussions involving NIWA, SAEON, George Percival (OGC), SN, and Jay Perlman (IEEE) to create a ‘BioData’ content specification for use in OGC services such as WxS and SOS.
      2. At Asilomar this was discussed in the WG 8 sessions, and agreed that TDWG and its liaison with OGC will be the channel for adoption.
   2. Sample Linkages
      1. GBIF (DH) wants to establish means of including unique identifiers to sample data into the DwC data stream.
      2. Concept has been documented.
   3. KNB (MetaCAT) mappings.
      1. Significant species observation data (as a minimum) are included into KNB MetaCAT repositories.
      2. SAEON has started mapping these data series as mediated Web Context Documents with a view to making them available as time series (simplified SensorWeb) or maps (GeoRSS has been considered).
      3. GBIF has a similar requirement (TR), but wants to map the data to an extended DwC schema.
      4. Probably best achieved using crowdsourcing (existing data) and MetaCAT-specific interfaces linked to IPT (future - see ‘Good Ideas’ discussion below).
   4. Consensus on these three programmes can be extended during the GBIF ‘hackathon’ 22-24 May.
2. Consensus document: content standards and initiatives to achieve this: WH to EoT - Scoping document for Workshop (end March)
   1. Input from DH – Sampling
   2. Input from TR – Mapping to KNB resources
   3. Input from JS and WH – inclusion into OGC services
3. Consensus document: process of adoption (TDWG, OGC): WH to EoT (end March)
   1. Process 1: Leading to proposals for TDW conference end of October, Florence.
   2. Process 2: Through TDWG-OGC liaison, propose to OGC at the soonest opportunity thereafter.
4. Both the consensus documents will be added as Annexures to the WG 8 Work Plan – working document can be found here:

# Book Chapter: GEO-BON Handbook

Discussion on this aimed to confirm the outline of the chapter. Significant progress was made. WH to summarise this in the WG 8 Work Plan document – latest version can be found here:

<http://saeos.dirisa.org/communities/geo-bon/reports/G372.2.2%20Framework%20-%20WG%208.docx>

# Some Good Ideas

1. WH to summarise concept in a note (end March).
2. Mediation (Annotation) of KNB data – links to content standard extensions to DwC in respect of KNB data, but probably goes beyond.
   1. Requirements for programmers to implement
      1. Draft – HS – Include KNB and EU-BON (mid-April)
      2. IPT examined– or metacat – for implementation
      3. IPT embedded into MetaCAT
      4. Denormalised flat files are required in some cases
      5. Maybe store denormalised data
      6. Mark Schildhauer(KNB) (MS) – HS to clarify questions.
   2. How do we normalize the GBIF data stream?
   3. EnvO – use as a preferred ontology?
   4. Use GBIF SKOS inputs and schema as references to mapped items?
3. Include into MetaCAT – derived from IPT codebase, EU-BON – HS – end 2013 – confirm milestones with SN (WH).
4. When specification is available – Web Context Document link? WH. (mid-May).
5. Allow users to contribute ‘mappings’.
6. Gathering meta-files into a registry –
   1. Vocabularies – <http://vocabularies.gbif.org> - not RDF
   2. EoT to confirm with Tim – mid-April.

# Summary of Tasks

|  |  |  |
| --- | --- | --- |
| Responsible | Description | Due Date |
| OM | Let others know when beta-testing services are available at <http://uat.gbif.org> | ASAP |
|  |  |  |