Document: 10 recommendations from DiSSCo

dhobern #1 April 17, 2020, 8:19am

1. Make the scope for the catalogue Natural Science Collections. This allows to include earth sciences collections (Geology, mineralogy palaeontology and anthropology collections) and life sciences (representing living organisms), and puts the focus on the most important characteristics: Natural and Science.
2. Within that scope the registry should be as inclusive as possible, and allow any collection unless it has clearly not a purpose for Natural Science, Natural History or Natural Heritage.
3. Make a separation between (1) institute collections, (2) private collections, and (3) other collections, and ensure that the institute is the primary source of information for an institute collection.
4. Make a separation between a registry of organisations and a registry of collections. Natural Science organisations are not different from other research organisations and do not need a dedicated organisation registry.
5. Provide a resolvable PID for collections through the Handle system in combination with a metadata profile defined for collections metadata to become FAIR (*Findable, Accessible, Interoperable, Re-useable*) (e.g. including the collection abbreviation, ROR, ORCID iDs and other relevant metadata).
6. Consider the feasibility to make collection abbreviations, widely used as labels for institute collections, globally unique.
7. Make the collection descriptions discoverable in the main search engines and other external services.
8. Make the collection descriptions machine actionable, and provide a DOIP (*Digital Object Interface Protocol*) interface.
9. Allow community contributions, roles with different authority, and provide clear ownership information and versioning.
10. Include attribution metadata through implementation of the TDWG-RDA attribution recommendation.

Document from Wouter Addink, Alex Hardisty, Sharif Islam, Claus Weiland, Falko Glöckner and Anton Güntsch for DiSSCo, with recommendations for the implementation of a global catalogue of natural history collections, April 2020 - available as **PDF**

1 Like

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

Faire Avancer le Catalogue des Collections d'Histoire Naturelle du Monde

建立《全球自然历史馆藏名录》

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

3.2. Community catalogues (TECHNOLOGY)

建立《全球自然历史馆藏名录》

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

Integrated summary from 17 to 30 April 2020

Integrated summary from 17 to 30 April 2020

Advancing the Catalogue of the World's Natural History Collections

---

dhobern #2 April 25, 2020, 12:04am

Many thanks to the DiSSCo team for putting this together. I fully agree with most of these recommendations, but here are some comments on recommendation 1 (scope of the catalogue).

I'd like us to think of nested sets of requirements and how we support each of these sets. We may end up with several hierarchically arranged catalogues that each have a purpose of their own and cannot properly support the more specialised versions. What I still don't know (and what I hope we can clarify) is which of these can be treated as "the same catalogue" and which will need to be kept separate:

1. A catalogue of **all scientific collections**. This is why GRSciColl was created. GBIF hosts this now and expects it to remain open for collections that are not related to biology or earth sciences. Whatever else we do, we need a mechanism to identify and reference *any* collection.
2. A catalogue of **all preserved biological collections**. This is a major driver for GBIF and many of the other stakeholders in this consultation. We have excellent use cases for this and we need to make sure we meet these needs. This needs to support and normalise much more detailed information than the catalogue of all scientific collections. The biological (species-oriented) focus is important for many of its uses. My main question is whether this can be built to include both of the next two classes of collection without weakening its effectiveness for these uses.
3. A catalogue of **all living biological collections**. Many of the uses of such a catalogue, and much of the content, will overlap with the catalogue of preserved biological collections. I'd like to know if there are any downsides we need to address before merging the two as a single catalogue.
4. A catalogue of **all geoscience collections**. These share a number of features with preserved biological collections and many institutions handle these resources as part of the same institutional collection. We need to think about how to meet these needs, but should we do so as a single catalogue or as separate catalogues (which could still be populated through a common pipeline based e.g. on a modular CD document format)? Most importantly, what are the use cases for a search or other data access that returns a mixture of preserved biological and geoscience collections? And are these use cases ones that would not be met by the catalogue of *all* scientific collections?

I think we should build all of these catalogues using the same infrastructure and tools and with modularised information in TDWG CD format. The issue is whether we need to brand these as separate catalogues with separate interfaces for different purposes. I am inclined to think we need to keep separate at least 1) The GRSciColl catalogue of all scientific collections, 2) A GBIF catalogue of all biodiversity collections, and 3) A geoscience collection catalogue (in partnership with **geosamples.org**?). All three would be built as a single information resource. In rough terms, GRSciColl would be the access point for all records, focussed on search through the generic sections of CD, and the other two would be based on all collections that provided modular CD data for the given type of collection.

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

Integrated summary from 17 to 30 April 2020

---

waddink #3 April 25, 2020, 7:01am

Hi Donald,
Where would0you see paleontology collections fit in these categories? Biological or geological?

---

dhobern #4 April 25, 2020, 7:51am

I think they probably fit into both - in a perfect world a modular CD format would include everything necessary to make these easily catalogued in both categories and hence to appear in “both catalogues”. I assume that geoscience researchers do search for paleontological materials as part of a geoscience search rather than seeing them as biological oddities outside their sphere …

---

pzermoglio #5 April 25, 2020, 5:24pm

Similar issues arise for ethnobotanical and zooarchaeological collections. I guess these would fall into **@dhobern** ’s 1. all scientific collections, and 2. all preserved biological collections. But given they too (as for paleo) have their own characteristics and requirements, should it be considered to include a 5. all anthropological collections? as mirroring the all geoscience collections. One could start with a few target catalogues in mind, but a nested structure that allows augmentation of general scope and leaves place for future interlinking as other communities get on board would seem a good way to go.

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

建立《全球自然历史馆藏名录》

Integrated summary from 17 to 30 April 2020

---

sabaj #6 April 27, 2020, 10:44pm

all living biological collections
Would this include zoos, aquaria and culture collections?

Based on my understanding, potential downsides here are the increased weight of complexity and the operational differences between institutions that curate living vs. dead biological collections (which promote a lack of connectivity between such institutions). Also, I think there are other registries in place for living collections. WAZA maintains a global network for zoos and aquaria. For microorganisms, NCBI's BioCollections manages a list of 820 "culture_collections". So, maybe the data is already out there.

Anyway…living vs. dead is one of the clearest distinctions one can make in the biological sciences (exceptions noted). I think No. 2, A catalogue of **all preserved biological collections** should be the initial focus and would be a **major** achievement in and of itself.

1 Like

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

Integrated summary from 17 to 30 April 2020

---

sabaj #7 April 28, 2020, 12:16am

Consider the feasibility to make collection abbreviations, widely used as labels for institute collections, globally unique.

Ideally, collection abbreviations would be globally unique. But, probably not worth the effort in practice as long as PIDs are used. For "Material Examined" sections, authors often copy and paste their key to abbreviations from one paper to the next. Those keys may not be up-to-date, referable to an established index, or vetted with the source collection. In short, it may be too difficult to get authors to consistently use a standard abbreviation. It might be easier to simply link multiple abbreviations to a single, stable PID.

1 Like

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

Integrated summary from 17 to 30 April 2020