3.2. Community catalogues (TECHNOLOGY)

dhobern #1 March 27, 2020, 10:35am

> This is topic 3.2. in the *Technology* section of the ***Advancing the Catalogue of the World's Natural History Collections*** consultation. Use this topic to discuss the questions listed below.

Background
IH is the best established catalogue servicing a large community of collections, but many other communities are important, including regionally or nationally focused efforts, such as CETAF's institutional profiles, the web portals of **https://www.idigbio.org/portal/collections[iDigBio]** and the ALA, and the **https://biss.pensoft.net/article/38772/**[One World Collection initiative], and thematically aligned efforts, such as the **World Directory of Culture Collections** and the **Global Genome Biodiversity Network portal**. A comprehensive global catalogue should ensure that the needs of these different communities are met and support their continued operation and independence wherever is valued by collections. Understanding these requirements is essential in planning the technical implementation and governance of the catalogue.

Other materials

The following contributed materials are particularly relevant to this topic:

- **Presentation: Index Herbariorum - Perspectives of a Long-Standing Collection Catalog**
- **Presentation: CETAF and DiSSCo Collections Registry**

Questions

- What catalogues already address the needs of some communities of collections?
- How can an integrated catalogue support these communities? Which communities require a separately branded identity and/or platform?
- What is the best way to include these communities as part of an interconnected solution?
- Is there a role for content to be created and improved by a wider audience (e.g. through Wikidata)?

Presentation: CETAF and DiSSCo Collections Registry

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

Advancing the Catalogue of the World's Natural History Collections

建立《全球自然历史馆藏名录》

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

Presentation: Index Herbariorum - Perspectives of a Long-Standing Collection Catalog

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》

Collections Catalogue - Daily Summaries

Integrated summary from 17 to 30 April 2020

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

Rich87 #4 April 22, 2020, 3:44am

One question that comes to mind after reading the intro to this topic:

How would one update their entry in a common global catalog if they are a member of a currently well-established community catalog? When GRBio first came into existence, herbaria were told they only needed to keep their IH entry current and the content would transfer to GRBio. Is there a similar thought being suggested in this catalog concept?

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

---

trobertson #5 April 22, 2020, 7:07am

Welcome to the forum @Rich87

> When GRBio first came into existence, herbaria were told they only needed to keep their IH entry current and the content would transfer to GRBio

The synchronization wasn't achieved between IH and the original GRBio implementation so GRBio became stale. However, working with IH we have *just* enabled this synchronization so this is indeed correct from April 2020 onwards. We are also going to migrate the iDigBio catalog data into GRSciColl (GRBio) shortly and then share the data management in a single database through the online administration interface.

> Is there a similar thought being suggested in this catalog concept?

Yes. Our thinking is that for each entry we should identify the authoritative source for the core information.

That could include for example:

- Using the IH record as the primary copy (or an alternative designated authority, such as a CETAF / DiSSCo registry)
- Using a metadata document associated with a dataset - many tell us this is the most accurate information
- Allowing an authorized editor to update directly in the catalog through web forms
- Working with collection management software developers so the entry is maintained directly

While this covers the core information (descriptions, staff etc) supplementary information such as citations, data use, indexes of specimens could then be linked to enhance the catalog. We have a sketch of how this could be displayed in an online catalog.

We'd welcome your thoughts on this approach.

---

dhobern #6 April 22, 2020, 7:39am

Thanks @Rich87 . I agree this is a key issue. In different situations, I can see institutions themselves maintaining their information in their own systems, or using something like Index Herbariorum as a registrar on their behalf, or perhaps getting assistance from the likes of iDigBio or the Atlas of Living Australia managing information at a national scale.

My view is that we need to map out a model for access permissions for managing and editing this information and that institutions should have the authority to claim their own records to be managed in whatever way they see fit. They should have the ability to delegate responsibility to IH, iDigBio, ALA, etc. or subsequently to reclaim the responsibility themselves. It may be that we would need a mechanism to ensure that institutions remained responsive and technically capable to maintain this information.

---

waddink #7 April 22, 2020, 8:09am

I think this authority issue mainly applies to institute collections (see our document for different profiles of collections we identified: Document: 10 recommendations from DiSSCo). I agree that the institutions should be the primary responsible for these records and that they should appoint an authoritative delegate like CETAF or IH to maintain the record in case they are not able to do so themselves. If this could be role based rather than system based (one or more people appointed by the institute to have that maintainer role), then it does not matter anymore in which system (CMS, CETAF registration system, GRSciColl) that person maintains the information (assuming that metadata like last modified date is stored and systems are synchonised).

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

Integrated summary from 17 to 30 April 2020

Maofang #8 April 25, 2020, 4:18pm

We have one question from @Wenjun posted in Chinese.

He wondered will the community catalogues adopt the same uniform standards as Integrated catalogue.

建立《全球自然历史馆藏名录》

Integrated summary from 17 to 30 April 2020

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

dhobern #9 April 26, 2020, 8:44am

Thank you **@Wenjun** and **@Maofang** . We expect that "community catalogues" will be subsets of the "integrated catalogue" and that the data standards will normally be exactly the same.

A community catalogue could just be a section of the integrated catalogue that is maintained actively by, or on behalf, of all collections from a country or for a particular taxonomic group. In this case, the standards, and the data, will be exactly the same for both purposes.

Or a community catalogue could be more like Index Herbariorum, with its own database managed by the relevant community. In this case, there *could* be some differences in the standards used, but we would aim to map the data for discovery through the integrated catalogue.

I expect that we will be flexible between these two models.

---

neilcobb #10 April 26, 2020, 3:20pm

We are developing an online catalog/database of all arthropod collections in the world **https://bug-collections.org** . It grew out of several NSF-ADBC Thematic Collections Networks created in the last 10 years (**Cobb et al, 2019**). It is an extension of our digitization efforts and is driven in part by the overall question- "Do we have enough specimens to address the full spectrum of key biodiversity issues?". There are scores of questions embedded in this general question about specific regions, time periods, and taxonomic and/or ecological groups.

We cannot answer this question unless we have basic information about collections, this is especially true for arthropods, since most of the data is based on specimens. Ultimately, we need to fully digitize all collections, but having a collections catalog will greatly help coordinate efforts.

To date we know there are ~223 arthropod collections in North America (Canada, Mexico, and the United States) that house 300 million specimens, not including lots and bulk samples. We estimate that collections are increasing holdings by 1% per year. Based on these estimates, we set a goal of 2,500 specimens per species for the ~170,000 North American species in order to adequately address the spectrum of questions from individual species to global fauna. This means that we need to increase our collecting effort by three-fold (i.e., 3% per year increase in holdings) to have enough specimen data unless we can coordinate more among collections. We cannot coordinate without having basic knowledge about collections.

We further estimate there are ~750 collections outside of North America, which means the current estimate for the total amount of arthropod specimens in arthropod collections in the world is one billion. If North American collections are representative and the estimates for number of extant arthropod species range from 1.5 to 7 million (**Stork, 2018**), then we currently only have 144-672 specimens per species. So even if we digitized them all we would fall far short of what we need. But these are still mostly fluffy estimates, we will know more when we complete the task of obtaining basic information about all collections.

Another result from our assessment of North American collections was that the even collections with more of a global focus (e.g., Smithsonian) were the most important collections in inventorying the regional arthropod fauna. So, having a small regional collection can be extremely important in providing data for an area at least within a radius of 100 km. The lack of physical collections typically means a region is poorly inventoried (e.g., northern Canada, northwestern Mexico, Nevada). Knowing where collections exist will allow us to either work towards creating new collections and/or coordinate efforts among existing collections to inventory under-sampled regions.

The work required to address the basic question posed above is monumental. For the World Index of Arthropod Collections to be effective in addressing biodiversity information needs it will be critical that the Catalogue of the World's Natural History Collections be immensely successful. It needs to define standards for concepts like "collections" and codes as well as provide an overall central repository for all natural history collections. It needs to establish protocols to seamlessly incorporate information from more specialized cataloguing efforts. There has been discussion already about working with herbaria at institutions with no arthropod collection to set up an insect cabinet so they could collect pollinators and herbivorous insects they find on their plant taxa of interest. Vertebrate collections already play a significant role in developing arthropod vertebrate parasite collections. So, it is not just knowing where arthropod collections exist that will help us obtain more arthropod specimens, knowing where all natural history collections are located will help.

There are other efforts that will greatly aid in building and defining the functionality of a Catalogue of the World's Natural History Collections. These include fishes (**A survey of digitized data from U.S. fish collections in the iDigBio data aggregator**), Mammals (**Mammal collections of the Western Hemisphere: a survey and directory of collections**), Mollusks (**Mobilizing Mollusks: Status Update on Mollusk Collections in the U.S.A. and Canada**).

2 Likes

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

建立《全球自然历史馆藏名录》

1.1. Directory to support the collections community (USE)

Integrated summary from 17 to 30 April 2020

---

gdadade #11 April 29, 2020, 1:09pm

> dhobern:
>
> What catalogues already address the needs of some communities of collections?

GGBN's registry addresses the needs of the (biodiversity) biobank community (currently 95 members)

> dhobern:
>
> How can an integrated catalogue support these communities? Which communities require a separately branded identity and/or platform?

GGBN will keep its branded identity and platform and is planning to retrieve its registry/catalogue data through the API of the new central catalogue. We will switch off our own registry and require to manage the biobank descriptive data in this central catalogue.

Collections Catalogue - Daily Summaries

Integrated summary from 17 to 30 April 2020

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

---

pzermoglio #12 April 30, 2020, 6:59am

From **@ErikaSalazar** in this **Spanish thread**

In Colombia there is the RNC, which maintains the inventory of the country's biological collections, and allows querying a directory of experts (the curators of the registered collections). You can access the information of each collection and review in more detail the biological groups held, the data about the collection holder, curator(s), types of preservation methods used, level of advance of taxonomic identification, cataloging, etc.

In the Colombian case, it would be ideal to have a separate record for the collections of preserved specimens and for those of living specimens that work as conservation centers exsitu (e.g. zoos, aquaria, botanic gardens).