2.1. Scope for the catalogue and definition of “collection” (INFORMATION)

dhobern #1 March 27, 2020, 10:14am

> This is topic 2.1. in the *Information* section of the ***Advancing the Catalogue of the World’s Natural History Collections*** consultation. Use this topic to discuss the questions listed below.

Background
The scope for the catalogue needs to be defined. The core use case under consideration is the listing and description of collections holding preserved biological specimens, referred to here as “natural history collections”. The consultation will focus on developing a solution that is robust and effective for natural history collections, but it is valuable to explore needs and opportunities around other types of natural science collection. Some of these may fit readily within the scope of the catalogue. In other cases, work on the catalogue may offer benefits to these other communities. Note in particular:

1. Many institutions hold both biological and geological collections and may manage these as a unified whole
2. DiSSCo includes geological collections within its scope, and other networks such as iDigBio include at least paleontological collections
3. GRSciColl was established to hold records on any scientific collection
4. The TDWG Collection Description standard is extensible for different collection types.

Other materials

The following contributed materials are particularly relevant to this topic:

- **Presentation: Global Registry of Scientific Collections**
- **Presentation: Index Herbariorum - Perspectives of a Long-Standing Collection Catalog**
- **Presentation: TDWG Collection Descriptions Data Standard Task Group**
- **Presentation: CETAF and DiSSCo Collections Registry**
- **Presentation: The Specifications of Earth Science Collections (CETAF)**

Questions

- What is the definition for our purposes (minimal and sufficient criteria) of a natural history collection?
- How do collections relate to and differ from institutions?
- How do collections relate to and differ from datasets?
- How do collections relate to and differ from collecting events (e.g. expeditions)?
- Should the following categories be included, or are there important linkages or opportunities that should still be considered?
  - Geological and paleontological collections
  - Anthropological collections
  - Ethnobotanical collections
  - Wood collections (xylaria)
  - Tissue banks, DNA repositories and slide collections
  - Living collections (microbial collections, zoos, aquaria, botanic gardens, seed banks)
  - Personal collections

Presentation: Global Registry of Scientific Collections

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

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

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

Presentation: CETAF and DiSSCo Collections Registry

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

2.2. Identifiers for collections (INFORMATION)

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

Collections Catalogue - Daily Summaries

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

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》

---

elyw #4 April 21, 2020, 4:05am

Grey areas for inclusion in this list include:

- archival material (particularly relating to historical collections) such as correspondence
- field note books
- collection registers including registers compiled by private collectors and are donated along with their collection.
- photographic collections, now also including video

2 Likes

建立《全球自然历史馆藏名录》

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

Integrated summary from 17 to 30 April 2020

---

dhobern #5 April 21, 2020, 5:39am

Thanks @elyw . You are correct. Topic 2.5 (Wider data linkages) identifies some of these grey areas as candidates for extended information that may be linked to a natural science collection, but from another perspective these things may themselves be seen as part of a specialised collection. We need to be flexible about how we handle the graph of interconnected information.

---

ekrimmel #6 April 21, 2020, 6:57pm

100% should include paleontological collections within the scope of this catalogue! GBIF is a primary distributor of specimen records from paleo collections, and many paleo collections are managed alongside extant biological collections in the same institution.

@tkarim

1 Like

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 #7 April 22, 2020, 3:01am

Thanks @ekrimmel . You are correct. Beyond stratigraphy and dating, and the problems we encounter mixing fossil and modern species, can you identify any needs that are unique to paleontological collections?

1 Like

---

waddink #8 April 23, 2020, 8:04pm

See also our answers to these questions in our document with 10 recommendations.

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

---

cboelling #9 April 24, 2020, 2:15pm

This is exactly the case in our institution (Museum für Naturkunde Berlin). Numerous artifacts like the mentioned correspondence, field notebooks and others exist and have on the one hand the role of information resources about the biological / geological specimens but are on the other hand administered as collections / aggregated resources in their own right. This is not only for administrative reasons (including storage needs). Many of these items are unique or have become so through, e.g., annotations of collectors or researchers which turn them into items with a museological value that contextualize the specimens they refer to in one way or another but are also informative for various fields such as history of science and various other cultural aspects.

Accepting a broad conception of what constitutes a collection in the context of the envisioned catalogue and the immediate connection these “grey areas” have to “proper” collections, I believe these could be generally included, possibly with content profiles of their own, as suggested in **this post**.

With regard to wether the catalogue should aim to represent deeper and more intricate relations among collections, such as above - it would be a good thing to choose a technological design which allows such representations to be added, possibly in a subsequent version of the catalogue, once the core structures are implemented.

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

Integrated summary from 17 to 30 April 2020

TuoHe #10 April 24, 2020, 10:55pm

I definitely suggest to include wood collections (xylaria) within the scope of the catalogue. Wood, as lignified plant, serves as quintessential materials used in most human history. The concept of wood collections fist appeared in the late 18th century, and large numbers of xylaria were established in the 1940s and 1950s. Institutional xylaria are listed in the Index Xylariorum, first published in 1967, and the most current edition is published **online** by International Wood Anatomists Association (IAWA) in 2016. It lists 158 institutional xylaria globally. A variety of national historical museums and herbaria also stores large number of wood collections, for example, Naturalis Biodiversity Center， Smithsonian Institution and Kew Garden. And large numbers of wood collections have been or being digitalized, databased and accessible online for public. Wood collections plays an important role in context of biodiversity and climate change, since tree rings could reveal and reflect the natural environment and climate in history. Reference wood collections are central to robust, scientifically sound forensic wood identification, which aims to combat illegal logging and associated trade to prevent biodiversity loss. The inclusion of wood collections in this list enrich the historical collections scientifically and technically.

1 Like

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

Integrated summary from 17 to 30 April 2020

---

dhobern #11 April 25, 2020, 12:13am

Many thanks **@TuoHe** . I fully agree. Xylaria are a very natural part of what we need to cover and meet all the criteria we expect for a preserved biological collection.

Are there any use cases you can suggest that would be unique to xylaria?

Anabela #12 April 25, 2020, 3:54am

> dhobern:
>
> Should the following categories be included, or are there important linkages or opportunities that should still be considered?
>
> - Geological and paleontological collections
> - Anthropological collections

> - Ethnobotanical collections
> - Wood collections (xylaria)
> - Tissue banks, DNA repositories and slide collections
> - Living collections (microbial collections, zoos, aquaria, botanic gardens, seed banks)
> - Personal collections

Well…I can see some issues with the Anthropological and Paleontological collections, because use to be sensitive in many occasions. In Argentina we have a really restricted law about those kind of collections. Also, some museums, like La Plata Museum (Argentina) has a restitution police for human remains who have been proven living offspring. But, even with all this, I strongly believe that all categories should be included, and also the grey categories mention by @elyw

1 Like

建立《全球自然历史馆藏名录》

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

Integrated summary from 17 to 30 April 2020

TuoHe #13 April 26, 2020, 8:08pm

Sure. Basically, xylaria provides valuable reference materials for forensic wood identification to combat illegal logging, which is a main cause of forest loss. Contrary to animals trees are stationary and show an open growth. Trees are subjected to environment influences during their life which may last up to millions of years, please refer to **Mahogany tree family dates back to last hurrah of the dinosaurs**. The xylaria could provide fossil wood collections for the tree of life and natural environment, and disk collections for research of the relationships between climate and tree rings patterns.

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

---

cboelling #14 April 27, 2020, 10:43am

What is the definition for our purposes (minimal and sufficient criteria) of a natural history collection?

I agree with @qgroom , @mswoodburn and others in **topic 2.2** that given the vast diversity of specimen types and perspectives on what is conceptionalized as an individual *collection* (possibly as a subset of the union of several other *collections* it may be hard to come up with necessary or sufficient criteria to define *natural history collection*. The (provisional?) definition from TDWG CD quoted by

> **2.2. Identifiers for collections (INFORMATION)**
>
> ‘a group of physical collection objects with one or more common characteristics’

could be a start both with regard to necessary and sufficient conditions . However, even this generic definition may run into problems. The Naturkundemuseum Berlin, for example, hosts the **Animal Sound Archive**. These recordings aren’t usually categorized as physical or material objects. They are information artifacts which can be copied from one physical bearer to another. Nonetheless, we regard it as an (important) natural history collection.

If I had to come up with a definition it would, provisionally, be this:

Natural history collection =def= A collection whose constituent parts (a) are derived from participant entities of natural processes and (b) have been collected to study properties of such processes.

The clause (b) is added to exclude cultural artifacts collections of e.g., wooden sculptures, stamps and historical weapons which are, at least in any conventional sense, not collected or designed to study the natural processes that brought their constituent parts into being.

This definition probably is as problematic as certain others. How important, in the context of the catalogue, is it really to have one? An informal description that conveys the intuitive understanding of “natural history collection” and intended use of the catalogue may be adequate, especially if the governance model allows for self-identification and self-governance of entries by participating institutions. It is, in my view, better to be more inclusive than restrictive with regard to grey area collections.

How do collections relate to and differ from institutions?

A *collection* is administered / held / owned by an *institution*.

1. An institution is an agent, usually a corporate body. Compared to collections, institutions belong to a materially distinct type of entity. Even if the holdings of an institution are identified as a single collection, the two entities should be represented separately.
2. At the whole-enterprise level, it is not uncommon to have the term “XYZ collection” denote the *institution* and at the same time the entirety of this institution’s collections.

3. An institution might administer a multitude of collections that are interrelated in various ways, notably by an overlap of the specimens accounted for in individual collections.
4. There might be cases where administration and ownership of a collection could be distinguished (specific use cases?). Who *administers* a given collection (and should be contacted about the collection) might, in the context of the catalogue, be more pertinent.

How do collections relate to and differ from datasets?

Collections are original entities, datasets are information about something. A dataset can be about a collection or some of its constituent parts. A collection can be represented in a dataset. In certain cases, the only manifestation of a collection may be in the form of information artifacts / datasets, e.g. digitally encoded animal sound recordings, i.e. without physical preparations.

How do collections relate to and differ from collecting events (e.g. expeditions)?

Collections are related to *collecting events* by virtue of the collection items which are part of the collection. Any collection item that is part of a given collection was collected - or is derived from one or more items that were collected - in the course of one or more collecting events.

Collections and *collecting events* are materially different types of entities. The latter is an event that has happened during some interval of time. The former is an object that keeps an identity over time (but may change over time with regard to its constituent parts).

1. In a variety of contexts the term “collection” is used to refer to collecting events or to collections as sets of items administered by an institution (each of which has been collected in a collecting event).
2. The *collecting event* from which an item originates may be unknown.
3. *Collecting events* may be nested or have other relations holding between them. An expedition, as a complex collecting event spanning months or years, may have numerous individual collecting events as parts (for which there might be detailed information pertaining to certain items in the collection relating to the expedition as overarching collecting event).
4. It might be challenging to actually define the *original* collecting event, especially when aiming to distinguish it from subsequent relocation or processing (like preservation or preparation). Intuitively, the collection event is the process in the history of a specimen for which there is no prior process of preservation or preparation. That initial event may be inseparable from processes that aim to preserve the material.

Should the following categories be included, or are there important linkages or opportunities that should still be considered?

- **Geological and paleontological collections**
- **Anthropological collections**
- **Ethnobotanical collections**
- **Wood collections (xylaria)**
- **Tissue banks, DNA repositories and slide collections**
- **Living collections (microbial collections, zoos, aquaria, botanic gardens, seed banks)**
- **Personal collections**

All of these fit the definition given above and should, in a more inclusive perspective on the catalogue advocated above, be eligible for inclusion.s

1 Like

Integrated summary from 17 to 30 April 2020

2.3. Hierarchical collection structures and subcollections (INFORMATION)

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

---

建立《全球自然历史馆藏名录》

2.3. Hierarchical collection structures and subcollections (INFORMATION)

---

sabaj #15 April 28, 2020, 11:43am

Because of their importance to taxonomy, I would add important historical/defunct natural history collections to your list, such as Museo Laurentii Theodori Gronovii, Lugdunum Batavorum [Leiden]. Also because some currently active collections are likely to become “historic” in the future.

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 #16 April 28, 2020, 12:22pm

Personal (aka private) collections. I think most taxonomists would agree that private collections should not be “legitimized” in any way. Yet, some taxonomists cite private collections in their works. And so, private collections cannot be ignored.

In some countries, the institutional infrastructure is too unstable…and so a taxonomist might assemble a personal collection in country. Such private collections are extremely valuable and warranted.

In countries with a robust institutional infrastructure, someone might assemble a private collection to control its use during their lifetime. Some personal collections masquerade as "museums" or occupy a professor's lab space in Dept X of University Y. Most of those collections eventually end up at an established institution of natural history where they often occupy "backlog" shelves and cabinets. Some are undoubtedly pitched.

Anyway…it is hard to both recognize and discourage private collections at the same time. I suppose private collections need to be recognized at some level because they are often cited in published works (many have established Codes!). Perhaps an ORCID ID would suffice as the PID for a personal collection.

---

neilcobb #17 April 28, 2020, 3:34pm

In corresponding with over 200 arthropod collections in North America as part of **SCAN** I am amazed at the variety of organizational workflows (or lack thereof) among the institutions. Generally, the institution does not seem to administer collections beyond asking how much funding they bring to the institution or how much they cost. Most US collections are based at universities and each collection is typically controlled by the curator. This is generally good, the institution relies on a tenure-track faculty to be successful. But it generally results in fragmentation. Collections typically have more loyalty to the CMS they use than sharing a common institution.

Institutions are generally permanent but the practices of individual collections are very idiosyncratic, changing with each succession in curators. Collections are much more likely to come and go. Most collections probably do not know what their institution code on GrSciColl is and several make up their own when the digitize.

Additionally, entomology collections in the US want to divide collections within a larger collection by taxa. For larger collections, especially at museums they already administer holdings by order or class.

I conclude that defining and tracking a collection is challenging because they vary so much among institutions, and they are not static, they will evolve, divide and combine with time as curators or institutional politics change.

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》

---

neilcobb #18 April 28, 2020, 4:19pm

I would also advocate for a new category "research collection". More and more we are seeing postdocs go for years without finding a permanent position and for the SCAN portal **https://scan-bugs.org/portal/** we are recognizing these as professional collections worthy of being treated with as much respect as institutional collections. We also post data from persons that curate their collections from their homes if they are actively collaborating with an institution or have plans to donate their collection to an institution and otherwise the curation of their specimens are done properly. I think the term personal collection is getting to be more and more dated.

---

waddink #19 April 28, 2020, 4:51pm

Hi **@neilcobb** , in DiSSCo linked projects we tend to call these Private collections (not personal collections, to indicate that these are different in ownership and access provision. Does that make sense? The distinction does not say anything about quality or size or value of the collection.

1 Like

---

neilcobb #20 April 28, 2020, 7:43pm

Hi **@waddink** , I am pretty sure that postdoctoral researchers would not like their collections to be considered private. This is an informal movement in the US that is just starting to gain momentum. I am just trying to promote digitization and GBIF will not accept data from personal collections. I think both personal and private have connotations of natural history stamp collectors. So we are trying to define a collection that wants to make data public and has high standards for curation. Also, when they do get a permanent job their research collection will go with them and be housed in a non-private facility. There are lots of private collectors that want their holdings private and I admit that I am only interested in people that want to digitize their collections and the only truly private collectors that I know are actively digitizing work with an institution and use their codes. I think "research" collections might be a small subset but I would not be surprised if there were at least 50 such research insect collections in the US.

waddink #21 April 28, 2020, 8:31pm

Thanks @neilcobb , this is very interesting. It is a class of collections I was not aware of. I am curious to know if such collections also exist outside the US. About negative connotations of English terms: in my experience this varies per part of the world, sometimes a term has a positive connotation in one part of the world and a negative connotation in another part. With multiple languages it probably becomes even more complex where in some cases you do not want to use the literal translation. But that is a relative small problem as users tend to get used to new terms quite easily as long as these are not confusing.

pzermoglio #22 April 28, 2020, 9:07pm

As a researcher, I would like to find the specimens that could inform my research. Those could be held in institutions or at someone's home (permanently or temporarily), and could have different access policies (loans policies from museums, or will to share from private collectors, etc.).
I think then there are three separate things here: 1) knowing that the specimens are there (somewhere), 2) knowing where they are physically, with as much level of detail as possible so that I can go knock on the door and ask for it), and 3) knowing under whose rules they are held, and which are those rules (who they depend on and if I'll actually be able to inspect them).
With this in mind, although we would all probably agree that the best situation is for specimens to be deposited in a place that allows open(ish) access (e.g., a public museum; understand place here as not only the physical spot but also the policies around it), I would want and/or need to at least know of the existence of those personal/private/other_denomination assets. Being personal/private or public may be subjected to change over time, as pointed out by others before, and passing from private hands to public hands may or may not ever happen (or not during the time of my research). Yet I want the data. And therefore I'd advocate for all kinds of collections (understood as for the preliminary definition of the CD group, *"a group of physical collection objects with one or more common characteristics"*) to be included.
I think it's a matter of thinking why do we want the catalogue, or rather, what for. Whether personal/private or else will be important to me only when I want to access the specimens. The question to me would not be whether to include the collections or not, but rather which information do I need to see attached to those records for those to be useful to me (like access options). What a person does with a collection they possess, e.g., if they finally deposit it or not, is a personal (ethical, if you let me) matter.

Yet, playing the devil's advocate, my first complaint would be in the lines of "but how do we trust this fellow declaring she/he's got a collection in her/his living room". So similar to the individual researcher's datasets not being directly published through GBIF.

- This is probably more of a governance matter, and not a scope one. I think we want those collections in, but we need a mechanism to ensure those are real collections. A system somewhat parallel to what GBIF uses to endorse organizations could be looked into, "nodes" and experts could be consulted about the collection in question (and this would only need to be done once at incorporation).
- But also an information matter, a minimum amount of info should be requested from these collections to be included (as for any). If for example an attribute of "is this openly accessible to the public", "where", "how", was to be declared, there would be no harm in trust from the community in the provenance of the catalogue records - maybe those holding back the specimens would even feel bad in comparison to others and actually open them :)

neilcobb #23 April 28, 2020, 9:53pm

Here is a real life example **M.Andrew Johnston Research Collection** .

1 Like

gdadade #24 April 29, 2020, 12:57pm

> dhobern:
>
> Should the following categories be included, or are there important linkages or opportunities that should still be considered?
>
> - Geological and paleontological collections
> - Anthropological collections
> - Ethnobotanical collections
> - Wood collections (xylaria)
> - Tissue banks, DNA repositories and slide collections
> - Living collections (microbial collections, zoos, aquaria, botanic gardens, seed banks)
> - Personal collections

A strong YES to include especially biobanks (meaning tissue, DNA and environmental samples), otherwise this catalogue is useless for molecular research at all, which today is one of the most important approaches in life sciences. GGBN already started this task many years ago and is urgently waiting for a central solution. These collections are usually overlooked, only the sequence and maybe the voucher specimen matters, which is a huge problem with respect to good scientific practice in general and the implementation of the Nagoya Protocol in particular. Traceability is crucial these days.

The living collections should be included, but maybe a synchronised approach like planned for IH is the way to go if they already have their own registries (I think at least the Zoos have something, but don't know for sure). Since GGBN already has members from all mentioned kind of living collections plus veterinary, crop and life stock collections we are happy to help working on this topic if needed.

1 Like

---

pzermoglio #25 April 29, 2020, 6:27pm

From **@maperalta** in this **Spanish thread**

1. Collections must have an institutional basis. Scientific publications should not include material from personal collections that are inaccessible to the rest of the community.
2. Data sets can involve distinct collections.
3. Collected material can end up in several collections from distinct institutions.

We should include geological and paleontological collections, anthropological collections, ethnobotanical collections, xylotheques, tissue banks and DNA repositories.

1 Like

---

bsterner #26 April 29, 2020, 10:08pm

Responding to these questions in the prompt: "What is the definition for our purposes (minimal and sufficient criteria) of a natural history collection? How do collections relate to and differ from institutions? How do collections relate to and differ from datasets? How do collections relate to and differ from collecting events (e.g. expeditions)?" I'm not convinced that the issue about what should be included in a general catalogue of collections is best settled by a definition of *what* a collection is. More likely, the question should be *which* collections are *appropriate* to include? Perhaps requiring validation of clear answers about legal ownership and public access rights may be more definitive. For example, would the catalogue include a personal collection whose owner illegally removed specimens from another country? Or what about a collection with no public access? Legal issues may provide a common ground here for determining minimal requirements that sidesteps the vague conceptual issue of what is a "collection."

WUlate #27 April 30, 2020, 1:09am

> dhobern:
>
> - How do collections relate to and differ from institutions?
> - How do collections relate to and differ from datasets?
> - How do collections relate to and differ from collecting events (e.g. expeditions)?

Let me attempt to answer with an example (apologies if I digress with content better related to other topics):
INBio (Costa Rica) had an Institutional Collection but the Research Department (Inventario) would probably clarify that it was actually several collections (Plantae, Insecta, Fungi, Mollusca, Nematoda, Arachnida, Myriapoda, Onychophora). Now Collection Managers (Curators) would argue there were different Collections ("Dry" Mounted Duplicates, Seeds, "Xyla", Wet, Vouchers Collections and so on) and each had its own protocols and management methods.
For example, a specialist could visit and take specimens of her group from the wet collection (of Malaise traps' soups) of insects, mount them and identified them for the dry collection. Some times, duplicates from a single collecting event would be sent to other institutions. Other times, live specimens will be grown in-house and by-products of the process would be catalogued as vouchers of Natural History associated to the specimen, the plant it fed from would go to one collection, the final dead specimen would go to the dry-mounted collection, next to all of its by-products.

Counting number of specimens was dependant on the definition of specimen itself that each Collection Manager decided (the mollusks in 1 rock, the shells in a vial, the sheets from the same plant, the nematodes mounted in one plate). Sometimes they will count each of them separately, sometimes they would count all of them together as one. It didn't matter what each collection considered an specimen, numbers were provided for the annual account and the reports to donors. In some cases, while digitization was still going on (not finished yet), an estimation was made using the same sampling method every year and the totals kept growing until, finally, the year the backlog digitization caught up, the total (not-estimated) number of specimens diminished considerably for some of the collections. Much later, INBio gave their collections away to another institution that finally incorporated them with their own collections.

Summarizing,

- There is a hierarchy in Collections that has to be reflected in the data handling.
- Collections will definitely overlap (thinking Dimensions in TDWG CD Standard here).
- A Collection could be housed in different Buildings/Departments but we are tending to an Institution-based model. So CollectionA[OneHalf@InstitutionX] is a different Collection than CollectionA[OtherHalf@InstitutionY] and a Collector's life-long collecting events (one possible dimension?) will also be divided by (Hosting) Institution, too.
- Depending on the Managers' definition of their Collections, a collecting event could produce specimens in different Collections.
- Total numbers estimated and reported by Collection Managers should maintain the calculation methods consistent throughout reports. Therefore, estimation method might be something to consider stating in the metadata (and maybe considering defining a categorization/vocabulary for it?)

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》

---

dhobern #28 April 30, 2020, 1:30am

Thanks @WUlate - this example is great. It certainly seems to reinforce the importance that others have already highlighted of letting institutions determine what breakdown into "collections" makes most sense to them, their staff and their users.

Your point about the same institutional holdings being treated both as a set of taxonomically organised collections and as a set of collections categorised by preservation methods is interesting. I could imagine an extreme case where a collection chooses to reorganise completely from the first model to the second (more likely than the other way around) and wishes/needs to present all of its holdings to the wider world as a new set of collections with complex historical relationships to the older ones (many partial transfers of old collections to new ones).

Ideally, in such a case, metadata would be clear enough that we would be able to determine both the original and the current collection for any specimen that preceded the reorganisation. In practice, I expect that we would need simply to handle the old collection codes as ambiguous synonyms for multiple new collection codes and rely on human effort or business logic to map older references (e.g. in an older taxonomic treatment) to the correct present-day one.

---

cboelling #29 April 30, 2020, 8:55am

I agree that the example by @WUlate nicely illustrates the complexities that may arise for the relations between collections in virtue of the histories of their constituent specimens.

Handling old collection codes as "*ambiguous synonyms for multiple new collection codes*" might fall short of achieving a meaningful representation of the actual relations among the collections concerned.

I would recommend drafting a carefully chosen and defined set of relations between collections that corresponds to user expectations in the context of the catalogue and that represents collections' relations along the lines of splits, merges, and in general transformations in terms of their constituent specimens (and I'd gladly try to contribute to such an approach).

This could be done so that both relations among collections existing in a temporal sequence (e.g., a historical collection being absorbed wholly or partially in a contemporary collection) and among collections existing concurrently (e.g., some specimens transferred from one to another in workflows such as the one described by @WUlate).

Last but not least such an approach could also support cases where collections, even concurrently, are conceived in terms of different partitions of the attribute space ("dimensions") of an institution's (or several institutions') holdings, e.g. one partition based on taxonomy vs. one based on preparation type.

1 Like

2.3. Hierarchical collection structures and subcollections (INFORMATION)

---

dhobern #30 April 30, 2020, 9:31am

Thanks @cboelling - I fully agree that we need the kind of approaches you describe and as far as possible to represent the actual relations among the collections. My "ambiguous synonymy" was because we know in advance that there will be cases where this is not possible and we need to be ready to handle the lowest common denominator as well as what we hope to be best practice.

---

waddink #31 April 30, 2020, 10:07am

I think a first priority for the catalog should be, given some of the potential usages mentioned in earlier discussions, to get a more complete and up to date overview of the institute collection holdings (the sum of the collections in an institute). Once that is established, AI tooling could be used to discover collection relationships where I think the most inportant use case is to relate mentions of collections in literature to current collections.

Adelantando el Catálogo de Colecciones de Historia Natural del Mundo

建立《全球自然历史馆藏名录》