

Annual Meeting of the US Virtual Herbarium Project (USVH), July 28, 2013

Location: Prince of Wales room/Riverside Hilton, New Orleans

Chairs: Mary Barkworth, Zack Murrell

Minutes recorded by Ellen Dean

Meeting started at 3:30 pm

I. Welcome

- Mary Barkworth welcomed the group. A sign-in sheet was handed around. Everyone introduced themselves by giving their name and institution.
- Barkworth went over what the USVH is – an Agricultural Experiment Station Project that grew out of the American Society of Plant Taxonomists (ASPT) Curators Meeting at Botany 2008. We are a coordinating committee with no funding. We promote the sharing of information between herbaria to facilitate digitization.

II. Activities of the USVH 2012-2013 and other announcements (Presented by Barkworth)

- A logo was adopted from those submitted in a 99Designs competition (Fig. 1) which is used on our Website.
- Eric Ribbens organized a symposium titled **“Herbarium Digitization for Research, Teaching, and the Public”** that would take place at Botany 2013 on July 29 at 8 am in the Elmwood Room. It is jointly sponsored by the US Virtual Herbarium project, iDigBio, the Society of Herbarium Curators, and the American Society of Plant Taxonomists. Financial support was provided by iDigBio (\$5000) and the Society for Herbarium Curators (\$300).
- Eric Ribbens created a new web site to replace that developed on the USGS site. It is located at <http://www.wvu.edu/usvirtualherbarium> and can be found by entering “US Virtual Herbarium” in a search engine.
- Barkworth announced that the 2013 survey of progress in digitization by US herbaria was being conducted using a Google survey form (see <https://docs.google.com/forms/d/1xcY-mT7r8VQ14io6wgu9FXEo1ZQc4GBGPsBI9Fdnpl/viewform>). Those present were asked to complete the survey for their own institution and encourage others to do likewise. By the time of the meeting, 75 responses had been received. The goal is to have 400 respondents.
- Results from the 2011 and 2012 surveys would be summarized at a presentation at 10.30 on Tuesday.
- Dean had inquired via the Herbarium Listserv about digitization funding sources. She summarized this information in a Word document which was sent out via the Herbarium Listserv and provided as a handout to those attending the meeting.
- Barkworth announced that there would be a joint meeting of the Society of Herbarium Curators and the ASPT Collections Committee on Monday, July 29 from 7-9 pm in the Belle Chasse room.



III. Network Integrated Biological Alliance (NIBA) presentation by Lucinda McDade

Lucinda McDade gave a presentation on NIBA. NIBA began as a series of workshops aimed at encouraging the National Science Foundation (NSF) to provide funding for the digitization of biological collections. This led to a strategic plan for NIBA which then led to establishment of the program for Advancing Digitization of Biological Collections (ADBC) by the National Science Foundation. This program has funded development of a central hub for all ADBC funded projects, iDigBio (see below), as well as establishment of several thematic Collection Networks (TCNs). McDade encouraged people to get on the iDigBio email list. The NIBA committee is now turning the strategic plan into an implementation plan which is available online (<http://www.aibs.org/public-policy/biocollections.html>). The American Institute of Biological Sciences oversaw creation of the strategic plan and is now providing oversight for its implementation plan.

The NIBA plan calls for three things:

- (1) Digitization of all US biological collections, large and small, using shared standards and formats, and making the resulting data web accessible.
- (2) Development of new Web interfaces, visualization and analysis tools, and georeferencing and data mining processes which will be made available for using and improving NIBA resources.
- (3) Enabling real-time upgrades of biological data and preventing the future occurrence of non-accessible collection data through the use of tools, training, and infrastructure.

The detailed recommendations in the implementation plan are geared toward attainment of six goals:

- (1) Establishment of an organizational and governance structure that will provide the national leadership and decision-making mechanism required to implement NIBA and to fully realize its Strategic Plan.
- (2) Development of the engineering needed to create an appropriate infrastructure for sharing information from US biological collections.
- (3) Enhancing the training of existing collections staff and creating the next generation of biodiversity information managers.
- (4) Increasing support for and participation in NIBA by the research community and a broad spectrum of stakeholders.
- (5) Establishing an enduring and sustainable knowledge base.
- (6) Infusing specimen-based learning and exploration into formal and informal education.

The NIBA process is now in silent phase while they try to move it forward. They may submit an NSF grant proposal for a Research Coordinating Network to move it forward. Implementation will need a lot of resources.

Discussion: Brent Mishler made the comment that herbaria also need resources to support the collections themselves, not just digitization money. Pam Soltis commented that in recent meetings with NSF, officers are encouraging them to make sure money is put into grant proposals for specimen vouchering and curation. Barkworth commented that the digitization wave is not sustainable if NSF funding is the only money available. Gerald (Stinger) Guala commented that the ADBC creation and NIBA process went forward in record time. If something else is needed for collections, we need a process and community buy-in. He commented that we need to write an influential report, because getting funding requires a process before the funds will be made available. Barbara Thiers commented that Judy Skog at NSF has said that we need to broaden the base of funding. It can't just be NSF driving these processes.

IV. iDigBio presentation by Pam Soltis

Pam Soltis gave a presentation on Integrated Digitized Biocollections (iDigBio) resources. iDigBio was established in 2011 and has five years of funding. There are 10 NSF-funded Thematic Collection Network grants (TCNs) – 3 funded in the first year, 4 in the second year, and 3 just funded. Plants are doing very well (Bryophytes/Lichens, a Tri-Trophic project that includes plants, and Macrofungi were funded, Macroalgae just funded, and several herbaria consortia have been funded). There had also been some Partners to Existing Networks (PEN) grants to herbaria or groups of herbaria.

iDigBio is trying to develop a portal and database that will house information for the ca. 1 billion US collections. The scale of this database is much bigger than ever before. They are now releasing version 2 of the portal. They are planning a release of version 3 in December or January. The data portal is at the iDigBio website. However, the database is not yet well populated, with only about 3 million specimen records. The Lichen and Bryophyte TCN has submitted the most information so far. iDigBio is developing mapping functions and links to the Taxonomic Names Resolution Service. They will also be including annotation modules so that people can comment on the names associated with records and images. She is very interested in feedback from different scientific communities, including the botanical community.

iDigBio is developing an Application Program Interface (API) – if herbaria consortia have data that they want to upload to iDigBio that are not generated by a TCN, this would be a way to access and serve those data. They hope to strengthen ties to international institutions too, so that they can share data.

They will be linking collections to genomics information and are developing an index to DNA banks in the US in connection with this goal.

They are trying to make the project community driven. Year 2 activities included lots of different training workshops on digitization and georeferencing. They have also cosponsored symposia, including two here at this meeting (unfortunately both scheduled for 8 am to noon on Monday). There are lots of upcoming workshops too. They are also working on some student and post-doc training in the area of digitization of collections. They have an exhibit table next to USVH table in the exhibit hall. She mentioned ways that people can interact with iDigBio. Some NSF-funded projects are required to give data to iDigBio. It is a contract in the award letter.

Discussion: McDade asked how this interfaces with the Global Biodiversity Information Facility (GBIF). Soltis said that they are working out this detail. They need to decide how the data will flow – possibly from iDigBio to GBIF to BISON (Biodiversity Information Serving Our Nation, a USGS program) or possibly via some other route. These entities will meet together next month. The data in GBIF and BISON are not equivalent to the data that iDigBio will serve. iDigBio is serving data from US institutions – and it is all collections, regardless of collection location. Currently BISON gets its data from GBIF. GBIF contains data from all institutions globally. BISON is serving all US-collected data from all institutions globally that provide data to GBIF. iDigBio will also be serving images and all sorts of information that come in from the different collections, not just the fields that GBIF provides.

There was a comment that there are faulty data in Genbank, and it is very difficult to annotate them as being wrong.

There was a question about how to integrate the Tree of Life project into iDigBio. Pam said that discussion has begun.

There was a question about whether or not data are checked before being served by iDigBio. Soltis responded that there is no editorial process. All that is checked is that data are in the correct fields. She noted that there is some quality control on the submitter end. Also, they will be building an annotation module that will send the annotation information back to the provider.

Question: As collections are annotated in a herbarium, how does that information get to iDigBio? How often are iDigBio data updated? Brent Mishler mentioned how important consortia are for gathering data from small institutions and that the aggregated data can be uploaded to iDigBio. A lot is happening locally that needs to be populated globally. Guala conceded that there is also a real time lag before data get to BISON. It has to be corrected locally, then goes to GBIF, then goes to BISON. Another comment was that, as soon as collections data are made available online, it greatly increases the number of requests for data, etc. Some commented that having data and images online had both decreased the number of inquiries they received and the amount of time required to respond to some of those they did receive.

V. Index Herbariorum (IH) update presented by Barbara Thiers

Barbara Thiers gave a presentation on the future of Index Herbariorum (IH). Maintaining IH has been a big time swamp for NYBG and has not been funded since the 1970s. She doesn't see anyone taking this on after she retires, and it needs to be self-updating. There is a plan underway to combine the Biodiversity Collections Index, the Biorepositories Index (developed by the Barcode of Life project), and Index Herbariorum. Thiers has agreed to do this so that people can update their own records. She has insisted that all the functionality of Index Herbariorum remain and the acronym (code) structure remain in the final product. She also wants to continue to moderate the entries – a lot of the foreign entries contributed by non-native English speakers need editing. The new product will be called the Global Repository of Biological Collections (GRBIO). The first iteration, which will not have all the functionality of IH, will go live in September. Users can provide feedback on this first version for all of this next year. The second version should be much more functional. For now, Thiers would like people to continue to send updates to their IH entries to her. Perhaps by summer 2014 there will be a fully functional GRBIO portal. The project is partly funded by Barcode of Life and Google. This portal is getting the eyes of the federal government – we need an index of all biological collections.

VI. Future of USVH discussion led by Zack Murrell

The US Virtual Herbarium project's term is up in 2014. Its goals currently overlap with those of NIBA, BISON, and iDigbio. The landscape of cyberinfrastructure changes all the time. Groups are formed and then are defunded. People are needed to move the information. We need to reach all the collections, both large and small. We need fresh faces and energy. We need input about what the role of the USVH should be going forward and how it can interface with NIBA, BISON, and iDigBio. The main role of the USVH is to help provide support and communication about digitizing collections. How can we move that process forward? Another problem is that there is at present no way to change the leadership of the project. Barkworth has said she will resign as co-Chair in September 2014.

Comment: Does USVH want to be a coordinating committee. Since it was an initiator, and now much of the work is proceeding with iDigBio, what is the current role of USVH? USVH needs to be more ambitious and tell others what needs to be done.

Murrell mentioned that the Herbarium listserv needs a home. The herbarium community is diffuse, and there isn't a governing body.

Barkworth mentioned that there are digitized collections that are not online. If that process can be made relatively simple, more collections can get online. She said that the USVH assists

smaller herbaria. Her surveys point out that smaller herbaria have trouble keeping up with the digitization wave, but it is easy for small herbaria to digitize using Symbiota, and is easier to get them databasing, because usually only one person has to make the decision to start. They can digitize a small herbarium quickly. Mary has seen people get going with very little funding.

Soltis: Last September, iDigBio gave a workshop for small herbaria on data entry options. They have another small herbarium workshop scheduled for this coming year. Sometimes small herbaria just need to get in touch with mentors in their area to show them how to get going.

Mishler commented that the US Virtual Herbarium project has not decided what its role is. In the past it has sometimes wanted to be a portal and sometimes a support organization. He thinks iDigBio is providing the support role, but the Portal is still needed. He thinks we still need a virtual herbarium like the Australian Virtual Herbarium.

Guala stated that we need to go forward with Mary's idea of teaching students and others not to leave legacy data for us to digitize. Students need to be taught to enter data online to make labels or create field notes in digital form. He emphasized that we need to get data into GBIF and thus to BISON. There would have to be a mapping tool at the portal in order to create an Atlas of Australia-like product. He said that there are products available via BISON.

Russell commented that the USVH has built visibility and helped coordinate herbaria and build community. The USVH website needs to emphasize the importance of how digitized specimen data can be used – we need to give those examples, such as Ackerly's work. We need to provide tools too. Herbaria are the envy of other disciplines; they are amazed at how far along the herbarium community is in terms of digitization, and it is partly because we have built a community.

Kevin Thiele (from PERTH, Australia, and one of the initiators of the Atlas of Living Australia): When the Atlas of Living Australia (ALA) came along, the Australian herbarium community had angst about whether the Australian Virtual Herbarium (AVH) was still needed; but they decided to build the Australian Virtual Herbarium on the back of the Living Atlas of Australia, and we could do the same by building the US Virtual Herbarium on the back of BISON and iDigBio. The Australian Virtual Herbarium also provides links to tools and ways for the herbarium community to keep in touch. Herbarium people tend to go to the AVH rather than ALA.

Guala stated BISON is supposed to be used the same way that the ALA is being used in Australia – as a source of data and mapping tools that all herbaria can access and use on their own websites or as a source of data for a national portal.

Harvey Ballard (Ohio University): The US Virtual Herbarium project is a way for people to come together and share experiences. It is a way to share information. People do feel it is valuable to share information and that the USVH project has been valuable in helping them do so.

The meeting adjourned at 5.30 pm.