

Minutes of the Annual Meeting of the US Virtual Herbarium (USVH) Project, 2011

Location: Maryland Room/Chase Park Plaza

Chairs: Mary Barkworth, Zack Murrell; **Minutes by:** Ellen Dean

Editors: Barkworth, Dean, Murrell, and Sellers. July 30, 2011 (Revised Sept. 6, 2011)

I. Welcome

Mary Barkworth welcomed workshop participants. Thirty-four participants listed their contact information on the USVH Workshop sign-in sheet.

II. Overview by Dr. Stella Coakley, Administrative Advisor

Dr. Coakley is Associate Dean at Oregon State University and Associate Director of the Oregon Agriculture Experiment Station. In the latter capacity, she is Administrative Advisor for various multistate Agricultural Experiment Station (AES) Projects based in the western U.S., including the US Virtual Herbarium (USVH) Project. She was previously Chair of the Department of Botany and Plant Pathology at Oregon State University, the home department of OSC, Oregon's largest herbarium. This experience gave her familiarity with herbaria. Multistate initiatives were formally instituted a number of years ago as a mechanism for coordinating activities among projects involving several AES stations. Some are research projects while others function as coordinating bodies. The US Virtual Herbarium Project falls in the latter category, being a five year coordinating project with an official starting date of 2009, making this the project's second year. For coordinating projects, individual Agricultural Experiment Station Directors within each state may approve travel funds for a state representative to attend the Project's annual meeting but such funding is at the discretion of the station director.

The US Virtual Herbarium Project is unusual in the number of its participants and the breadth of their affiliation: 68 listed individual participants representing 15 land grant universities and over 40 other organizations. The "other" organizations include non-land grant universities, a few free-standing museums or herbaria, and representatives of the US Geological Survey (USGS).

III. Review of 2010 meeting and progress on project goals – Mary Barkworth

Barkworth presented the Mission Statement that was developed for the project. This is available on the Project Web site www.usvirtualherbarium.org.

The Project seeks to involve all herbaria nationwide and recognizes that i) All herbaria have unique collections that provide information; ii) The next generation of students need to know the current standards for collecting, label-making, and herbarium curation, and many of these students are at small colleges; and iii) reaching all institutions will help foster an interest in biodiversity and collections.

Barkworth stated that it is her impression that the Project has broadened the level of interest in collections and herbaria but that it is difficult to separate its impact from that of general interest in making information more widely available and recent changes in National Science Foundation (NSF) funding.

The USVH Steering Committee defined three goals to be reached during the period October 1, 2010 – September 30, 2011:

Goal 1. Obtain information on the current status of all herbaria in the U.S.

Responsibility – Mary Barkworth.

Progress: The information needed by the US Virtual Herbarium Project includes a listing of the extant herbaria in the US, the number of specimens they hold, and the number of specimens that have been databased and imaged. Working from information provided in 2009 by Dr. Barbara Thiers, who maintains *Index herbariorum* (IH), Barkworth revised the list of extant US herbaria, eliminating those whose collections had been transferred, and adding the new herbaria reported by IH on Twitter (<https://twitter.com/#!/ihupdates>). She then georeferenced the extant herbaria, using Google Earth, and sent a request for updated information on holdings, databasing, imaging, and web-based impact to the individual named in IH as the contact person. The

requests were sent by Email if an email address was listed in IH, by regular mail if an email address was not available.

Slightly more than six hundred enquiries were distributed. There were 229 responses received as of 1 July 2011. Information from several herbaria in the SouthEast Regional Network of Expertise and Collections (SERNEC) were obtained directly from its Web site. Seven mail inquiries were returned as undeliverable. About 10 herbaria on the extant list were reported as closed. Reasons for the closure of herbaria varied. Sometimes a program had been discontinued; sometimes the person in charge retired and the next person was not interested; in one instance it was lack of interest by anyone outside the institution. In only two or three cases was the herbarium known to have been transferred.

Figure 1 shows the distribution of herbaria in the US. Larger dots represent responding herbaria. Among the responding herbaria, there were 43,701,327 specimens, of which 14,617,125 (33.4%) had been databased and 1,384,941 (3.17%) imaged. These numbers are impressive. They are probably biased in favor of herbaria that are engaged in digitization. Nevertheless, they demonstrate the commitment of those in charge of herbaria to databasing their collections.

The size distribution of responding herbaria was very similar to that of all US herbaria reported in IH, suggesting that interest in digitization is truly pervasive within the US herbarium community (Figure 2).

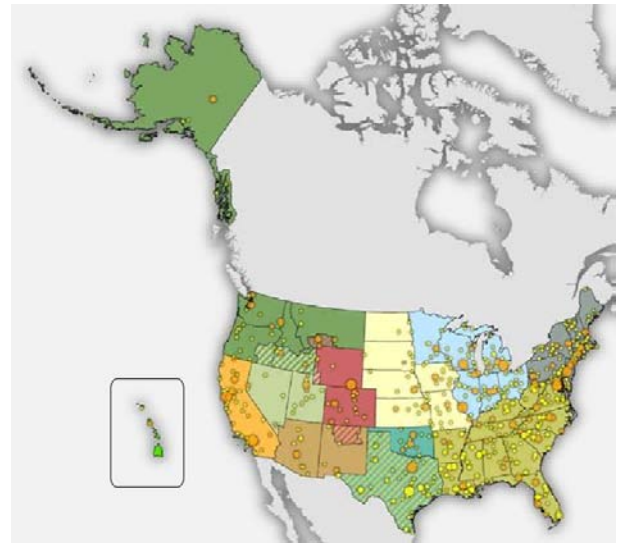


Fig. 1. Geographic distribution of herbaria in the U.S. Larger dots represent responding herbaria.

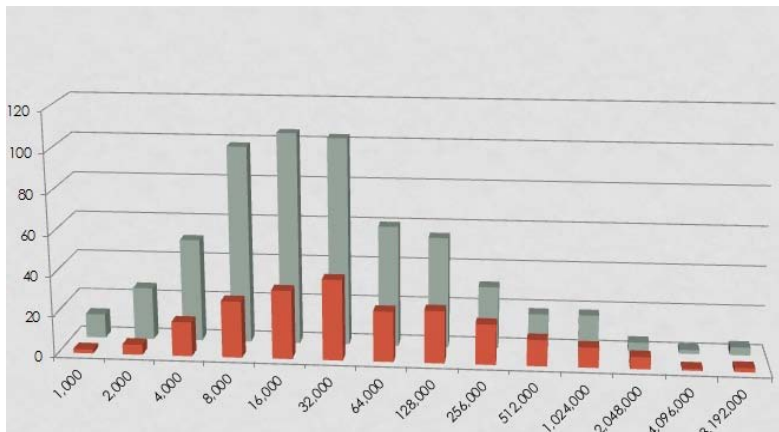


Fig. 2. Comparison by size of responding herbaria with all US herbaria. X axis, herbarium size; y axis, number of herbaria. Gray – 2009 data for all herbaria (Source: IH); red – 2011 data for responding herbaria.

Most databases were accessible only within the herbarium concerned. Of those that provide online access, most do so through their own server. There is a need for establishment of more regional nodes so that locating information and databasing itself can be conducted more efficiently.

Independently of the US Virtual Herbarium survey, Thiers arranged for georeferencing of all the 6,000+ herbaria in IH. This has resulted in her receiving an unprecedented number of update requests.

Next steps: In mid-August, Barkworth will attempt to obtain information from the approximately 350 herbaria that have not responded. In September, Thiers and Barkworth will discuss how the US Virtual Herbarium Project can assist Thiers in maintaining the currency of information in IH for US herbaria without increasing the work of maintaining IH. The desire of the US Virtual Herbarium Project is to be able to show annual progress in the digitization of US herbaria.

Goal 2. Develop three presentations concerning digitization.
Responsibility – Mary Barkworth.

Several individuals, particularly individuals at smaller herbaria, have asked for information on how to make progress in digitization. In response to such requests, Barkworth has started to develop presentations on a) barcoding specimens, b) making herbarium specimens, and c) information that should be on a label. In addition, Ben Legler (WTU) is looking into the preparation of a video on imaging vascular plant specimens. These presentations are designed to provide an example of effective practices. They incorporate a mix of "slides", sound, and video clips and will be web-accessible. Preliminary versions of three presentations will be available by September 30, 2011.

Goal 3. Develop funding of about \$15,000 per year for administration of the project. Responsibility – Mary Barkworth.

There are various costs associated with the US Virtual Herbarium Project, e.g., payment for a table in the exhibit hall, student assistance in conducting the survey, registration/ownership of the usvirtualherbarium.org and .net domains. Moreover, it would be desirable for the project to be represented at the annual meeting of the Society for Preservation of Natural History Collections (SPNHC) at least every other year.

No progress has been made on this goal, in part because it would be easier to approach potential donors when we have more reliable information on the number of herbaria and specimens in the US. Barkworth is working toward creation of a non-profit organization, tentatively called Promoting U.S. Herbaria (PUSH). Such an organization might also make it easier to funnel money to multiple herbaria with lower (or no) overhead than is currently possible.

IV. Synopsis of activities since the 2010 meeting, including report on current status of herbarium digitization (Zack Murrell).

In February, 2010, there was an invitation-only meeting at the Missouri Botanical Garden to identify the needs that should be addressed in order to achieve the goals of the US Virtual Herbarium Project. The 64 attendees included experts in information technology and bioinformatics as well as taxonomists. They agreed that community-building is important. There needs to be a support network connecting those that have resources and active collections with smaller herbaria.

A. Reports from Taskforces were that were formed in February.

The taskforce compositions and charges are available on the Web site www.usvirtualherbarium.org/MULTISITE/USVH/?q=taskforces.

Taskforce 5. Communications Infrastructure. (Leader: Elizabeth Sellers, USGS Biological Informatics Program). Sellers and other members of the taskforce have developed a Web site for the project (www.usvirtualherbarium.org) with Sellers as Content Manager. The Web site was built using Drupal to provide a Web presence for the Project and to provide online collaboration capabilities for the taskforces and other user groups such as regional networks. It seems, however, that most such groups are using other communication mechanisms so the online collaboration functionality of Drupal is not needed. This being the case, she is likely to move the project Web site to a format that is easier to maintain (e.g. HTML). Sellers is also working toward migrating the Herbarium listserv (<https://www.nacse.org/mailman/listinfo/herbaria>) (currently hosted by the Northwest Alliance for Computational Science & Engineering (NACSE) and managed by Dr. Aaron Liston of Oregon State University) to a USGS site because Liston has stated he would like to give up responsibility for the listserv. Although it is an open, moderated listserv, it has been the primary means of communication used by the Project. Liston has not received any complaints about its use in this regard. The only obstacle to completing its transfer that remains is providing access to and continuity of the listserv archives.

Taskforce 8. Integration Model Development. (Leader: Michael Denslow)

Taskforce members have been investigating how to create a portal to serve integrated data. They got together before the February meeting to discuss tools that they could use. They have developed an initial collaboration with Lifemapper (www.lifemapper.org/) at the University of Kansas Biodiversity Institute, which has a niche-modeling geospatial development tool.

Another question has been how data for the US Virtual Herbarium should be harvested. Initially, it will be simplest to obtain data from regional herbarium networks (www.usvirtualherbarium.org/MULTISITE/USVH/?q=content/regional-networks). The US Virtual Herbarium Project is committed to building on the work of regional herbarium networks and state-level consortia but it is also desirable to enable integration at a national level. Development of such integration will require financial support.

The task force also recommends that the project develop a “best practices” document or series of documents on herbarium digitization.

Janet Sullivan (NHA) asked if the Project or any of the regional consortia had a data sharing agreement in force. Barkworth replied that the US Virtual Herbarium Project itself did not have one as yet. They were working on the Global Biodiversity Information Facility model – if we had already agreed to provide data to GBIF, then no need for another agreement.

Taskforce 9. Professional Development (Leader: Patrick Sweeney). The taskforce recommended that the Project sponsor a workshop at the next SPNHC meeting which will be held in June, 2012 at Yale University and that it develop a list of the biodiversity and informatics specialists in each region.

Taskforce 10. Governance Development. (Leader: Mary Barkworth). A governance structure for the US Virtual Herbarium project has been discussed, amended, and the amended document was approved by the Executive Council in January, 2011. She recommended that the taskforces be disbanded. In further discussion, she commented that the amended governance document states that the project participants must take a vote in the project’s fourth year on whether it should be continued. This provision was included in response to a comment that there was no need for another society to speak in support of herbaria; there was support for a project to support digitization of herbaria.

Zach Murrell pointed out that there are multiple entities/vehicles for moving the community forward. The recent HUB award may supplant many of our efforts.

Reports were not received from:

Taskforce 1. Collections metadata aggregation (Leader: James Macklin).

Taskforce 2. Concept mapping and Name Resolution (Leader: Alan Weakley).

Taskforce 3. Funding and Partnerships (Leader: Brent Mishler).

Taskforce 4. Public Outreach and Education (Leader: Lucille McCook).

Taskforce 6. Scientific outreach (Leader: David Ackerly).

Taskforce 7. Digitization clearinghouse (Leader: Bil Alverson).

B. Reports from Regional Consortia

Note: All the regional reports will be posted to the project’s Web site. The following is a synopsis of comments made at the meeting.

Alaska (Steffi Ickert-Bond via Barkworth). The report showed that almost all the herbaria are in the process of digitizing their collections, some being engaged in discussions as to how to do it, others being engaged in the process.

California (Ellen Dean): The Consortium of California Herbaria received funding in September 2010 from NSF-BRC (ca \$1.9 M over five years). The grant is focused on databasing plants that will be most affected by climate change and georeferencing specimens. The Principal Investigator institutions are California Academy of Sciences, Rancho Santa Ana Botanic Garden (RSA), UC Berkeley, UC Santa Barbara, and UC Davis, but money is being distributed to other smaller institutions through RSA, Santa Barbara, and Berkeley. About 19 primary data providers are involved in the Consortium at the present time and the total number of specimens already available at the Consortium Web site is just under 1.3 million. Approximately half of those specimens are georeferenced. The Consortium focuses on serving data from specimens collected in California, and the Consortium grant provides funding to integrate data from California specimens at Harvard University and by

linking to SEINet. In addition to this joint grant, some California herbaria were awarded other BRC grants in the past two years. Those awards are described in more detail in the detailed report that was submitted to Barkworth.

Great Plains. No report presented.

Intermountain Region (Mary Barkworth). Herbaria in southwestern Idaho have received NSF-BRC funding. Many small herbaria are involved. Utah Valley University, which is fully databased, continues to image and database all of its incoming specimens. A regional portal has been established using SYMBIOTA software (www.intermountainbiota.org). None of the major herbaria in the Intermountain Region have received funding but they are continuing their databasing efforts. SYMBIOTA software permits creation of teaching checklists that include images of living plants. These have been created for the Plant Identification competition of the Society for Range Management and various Future Farmers of America (FFA) plant lists for Utah. These have been greeted with enthusiasm by those responsible for coaching students involved in these competitions.

Northeast. Janet Sullivan gave the report for Patrick Sweeney (Chair). The Consortium of NE Herbaria met in Philadelphia this past year. Forty people, representing 9 states and 3 Canadian provinces, were present. There were field trips on Sunday and talks on Monday. On Tuesday, there was a SYMBIOTA software workshop. Visitors from outside the Consortium were invited and there was a Global Plants Initiative speaker.

The common goal of those attending was to unite all the herbaria of the region in providing online access to data. Everyone shared information. They are moving ahead despite an absence of Consortium funding. Some member institutions are using grant funds or internal funds to move forward. Over 1 million specimens have been databased. The long-term goal is to serve these data over a regional portal. Those present beta-tested SYMBIOTA (<http://symbiota.org/tiki/tiki-index.php>) by providing datasets from each institution. Currently, 200,000 records are available in the SYMBIOTA database. Beta-testing is ongoing. They are working on data-sharing agreements and discussing the concept of protecting sensitive plant data. They expect that their portal will go live in late fall. Some members are waiting until this happens before making their data available to it. Currently, the Consortium's major goal is to establish the portal and to help students use it in preparing floras. A short discussion ensued on this topic.

POST-MEETING ADDENDUM from Barkworth: Further examination of the GBIF agreement suggests that a different agreement is needed for members of the US Virtual Herbarium Project. The GBIF agreement does not permit any entity other than a national node to provide access to its records for analytical purposes. In other words, under a GBIF-like agreement, regional networks could not make records from their contributors available for use by other portals.

Pacific Northwest. Ben Legler. The Pacific Northwest region is in the first year of a three year Consortium grant. June of 2010 was the start date. The Consortium includes British Columbia and the northwestern states (Alaska, Idaho, Montana, Oregon, and Washington) and the Yukon Territory. They have bought imaging equipment (digital SLR cameras, and light boxes) and delivered these to smaller herbaria. Students are hired at those locations to image the specimens using the equipment. About 150,000 images have been recorded, mostly from smaller herbaria. The databasing is being developed from images captured by undergraduates at the larger herbaria (REU students this summer). The reason for hiring students from the larger herbaria is that employing students from the smaller herbaria would incur overhead charges from both a PI's institution and that of the small herbarium. Some optical character recognition (OCR) technology is being used in databasing. Almost all the bryophytes at the University of Washington and Oregon State University have been databased. New databases are being set up for some herbaria. Herbaria in southwestern Idaho have received their own BRC funding on its own and are using it to participate in the PNW project. When the portal goes up, they anticipate having ca. 1.5 million records.

Pacific Islands. Michael Thomas. The Pacific Island Consortium received a BRC grant this past cycle. There are three US institutions involved: the University of Hawaii, the Bishop Museum, and the National Tropical Botanical Garden. They are two to three months into the process which will see the digitization of approximately 1.3 million records when finished. The three US institutions are partnering with institutions in Fiji,

Samoa, Guam, and other Pacific countries. Buying equipment for training programs at these places has begun. The Bishop Museum is the largest herbarium involved, having 0.5 million records to database. Some of the international cooperators have concerns over how the data will be used. Traditional knowledge is often recorded on specimen labels, and there are concerns about sharing these data. Although the specimens themselves are already available to the public, this project will facilitate international/increased access to the information, and there are concerns about the consequences of making such widespread access available.

Southeast. Zack Murrell. SERNEC developed a complete list of the 223 herbaria in the region. About 150 are actively curating or collecting. SERNEC held a workshop in Georgia . They also held a separate Algae collections workshop. They are working with zoological collections in the area and have helped conduct a survey of all collections in the region. There are about 200 identified collections and many are not part of any index (see list at www.serneec.org). SERNEC also offered a workshop on georeferencing in April, 2011 and are scheduling another three day workshop for October, 2011. They affirmed their desire to form a SE chapter of the Society of Herbarium Curators. Other initiatives include Mellon Foundation support for imaging Type Specimens. Other projects include one focusing on the Gulf Primary Producers and one based at UNC-Wilmington that involves an all-taxa digitizing effort.

Southern Rocky Mountains. Ron Hartman via Barkworth. The consortium received NSF funding in 2003 to database specimens from Colorado, northern New Mexico, and southern Wyoming. Funding has also been provided by home institutions and federal agencies. A significant percentage of collections from the region have been databased. Colorado State University received NSF funding this last cycle to image 97,000 specimens over the next three years. The Rocky Mountain herbarium has many collaborations that have led to funding being provided for databasing or imaging. Data from the consortium are served through individual member institutions, SEINet, IntermountainBiota.net, and GBIF.

SEINet. Liz Makings. SYMBIOTA is being used to integrate numerous datasets. The SEINet instance now has: a search engine, distribution maps, image library and field photos, regional checklists, and interactive keys. This collaboration originally began as funding for Arizona herbaria. Now 27 institutions in five states and Sonora, Mexico contribute data. They have done some workshops, including one for TORCH. They recently added three smaller institutions. Live data management through a web portal has been implemented. This enables members to make corrections, add georeferencing, etc. directly into the database. This editing is live at Arizona State University (ASU) and some other places. ASU received a grant to digitize Latin American specimens (including imaging); they are using OCR software to capture data and are about 60% done. They are developing three-dimensional distribution maps. Their model has been used by some other groups in South America.

TORCH (Texas and Oklahoma). No report presented.

Western Great Lakes. Harvey Ballard. An NSF-ABI grant proposal was not funded so the Consortium is investigating Institute of Museum and Library Services (IMLS) grant funding. Several smaller initiatives are going forward. On the other hand, Ohio's NSF proposal was successful. The institutions involved have completed databasing about 70% of their collections, including all Ohio holdings. The institutions involved are now switching to imaging two Ohio institutions that are already databased. Twelve herbaria have databases and the 13th is using Excel. Ballard has been helping the smaller Ohio herbaria get started. The next task is to get a network established. Ballard is thinking about how to obtain small amounts of money to do some "blitz" databasing at smaller herbaria using his undergrad students.

C. Update on the USGS and NBII. Elizabeth Sellers (USGS). Sellers has worked on building the National Biological Information Infrastructure (NBII) since 2004. The NBII was coordinated by the USGS Biological Informatics Program. As a line item in the President's Budget, the NBII was scheduled for termination in FY2012 (Ref: Terminations, Reductions, and Savings, Pp. 48 www.whitehouse.gov/sites/default/files/omb/budget/fy2012/assets/trs.pdf). The US Geological Survey itself has also received major budget cuts, which have forced the agency to initiate early termination of the NBII Program in FY2011 (rather than in FY2012). However, the USGS Biological Informatics Program remains as a component of the USGS Core Science Systems mission area that was formed as a result of the recent

realignment of the agency. Sellers and the USGS Biological Informatics Program staff are currently working on evaluating and migrating NBII-hosted datasets and other information products to alternate environments. Users will start to see rebranding of Web interfaces and changes in the Web domains (e.g. from NBII to USGS) occurring over the coming months.

The USGS Biological Informatics Program has also initiated a data integration project to make USGS biological and other datasets more accessible and interactive. The 'Biological Information System of the Nation' or BISON will build on and also serve as the US node to GBIF and make GBIF-served data more integrated with other federal and non-federal datasets. We hope that the USVH will be a major contributor and user community of the tools and services that will be offered as part of the BISON project.

V. Interaction between the USVH project and the HUB. Nico Cellinese. The NSF award for organizing the HUB (Home Uniting Biocollections) was made to a group led by the University of Florida. It has named the HUB "iDigBio". It will provide a framework for digitization of US biological collection datasets and will work on tools to digitize specimens at a greatly accelerated rate. In addition, three Thematic Collections Networks (TNC's) were funded. These involved 92 institutions, including at least 20 herbaria, across 45 countries.

iDigBio's role is, in part, to connect all the dots – serve all the data from all the different national collections. The US Virtual Herbarium Project participants may provide the legacy plant data. USVH can use iDigBio as a forum, discussion blog, to form goals, increase communication.

A presentation will be made on "Digitization Methods, Technologies, and Standards for Biodiversity Collections" at the 2011 Taxonomic Databases Working Group (TDWG – now known as 'Biodiversity Information Standards') meeting, October 16 to 21 in New Orleans, LA. Cellinese urged those present to attend if they could.

More information about iDigBio was distributed during the Curator's meeting at 7 pm Monday. Judy Skog was there to talk about the changes in BRC and Systematics NSF programs. There was also an organized brown bag lunch meeting at which NSF representatives provided information on changes going on within NSF.

VI. Identification of goals for 2011-2012

Develop a nomenclatural information resource – a Web site that provides nomenclature and quick access to information about the reasons behind nomenclatural changes in US plants. Suggested by Barkworth. Existing resources such as TROPICOS provide information about how different works interpret plant names but do not provide information about why changes have been proposed, nor a clear statement as to whether a change has been widely accepted or not. Barkworth suggested that having a place that provided access to such information would be helpful and, if the task were split up, would not be onerous to develop particularly if the FNA volumes became, as published, the starting point for discussions.

Comments made:

The Global Plant Checklist (MO and K) has been investigating this issue. People want the reasons behind name changes. Transparency is important. It is important to know how names relate to each other across state boundaries.

Wikispecies allows people to submit comments and update and edit. Most lists online do not allow those comments. Perhaps this would be a useful tool to start with.

Inform iDigBio about our work to date. Herbaria wish to be an integral part of the iDigBio project. We need to inform the herbarium community about best practices so that the data is useful by the HUB and to inform iDigBio about our needs and problems. In addition, we need to determine how the US Virtual Herbarium Project can most effectively interact with iDigBio.

Integrated access to species information. At present, existing nodes provide access to data from their member institutions. Integrated access means that any portal should be able to provide access to all data

being provided by herbaria. One approach would be to address this through the nodes. There wasn't much enthusiasm for this idea at the meeting partly because the idea of a US national data portal was controversial. At the February meeting, it was not clear that there should be a national portal. Questions that were raised included whether data should be provided to iDigBio or the US National node run by USGS, or GBIF.

Nomenclatural Crosswalks. Integrating data faces the problem of harmonizing differing taxonomic treatments. A thesaurus can serve for a limited region but as the number of institutions involved increases, the task becomes more complex.

Comments made:

The Consortium of Northeastern Herbaria (www.neherbaria.org/) portal will allow people to choose the taxonomic system to use when searching, e.g., FNA, GRIN, etc.; their portal will allow people to retrieve all the specimens relating to a name within a given taxonomic synthesis.

Other initiatives that may be addressing this issue include iDigBio, BiSciCol, and the taxonomic names resolution service of IPlants.

Communication. There was a strong request that there be more frequent and more timely communication concerning the US Virtual Herbarium Project. For instance, the annual report and agenda should be distributed several weeks before the annual meeting.

There should be a follow up to the annual meeting that lists the action items. It should be distributed within 30 days so as not to lose impact.

Coakley emphasized the importance of reporting on the impact of our activities within each region. Address the question of why does this project matter? Make the project better known. **Everyone involved needs to contribute in this respect.** We also need to make sure our professional societies (ASPT, BSA, SPNHC, NSCA, etc.) are aware of what we are doing.

Sellers stated that she is willing to post information to the USVH Web site at any time.

By next year, each consortium should come to this meeting with a list of herbaria in their region and the contact information for those herbaria. It was also suggested that the project create a database of projects and expertise. Barkworth requested that each Consortium include at least one activity involving engagement with K-12 teachers in the coming year.

Action items and target dates

The following items are listed by their target dates, not in order of importance.

- Post report on annual meeting and reports from regions to the Web site, Herbarium listserve, and NIMSS. August 13.
- Complete and post results from survey. September 30.
- Complete version one of three online presentations relating to digitization. September 30.
- Formally register a fund-raising entity for promoting US Herbaria. September 30.
- Develop an updated directory of contacts and herbaria in each region. October 30.
- Form a small committee to investigate effective mechanism for creating an informative nomenclatural resource. Preliminary report to Steering Committee January 31, 2012.

Comments

Please submit comments/feedback regarding these minutes to the US Virtual Herbarium Project Chairs: Mary Barkworth (Mary.Barkworth@usu.edu) and Zack Murrell (murrellze@appstate.edu)