The Case for Collections Support: Getting to the ‘First Step’

(Title) Austin, in his abstract for this symposium, has noted that as a ‘first step’, NSF has put funds into a program on Advancing Digitization of Biological Collections. He then asked me to provide some background information similar to the presentation I gave last year at the Biodiversity Informatics Standards (TDWG) meeting. I would like to emphasize that it has taken a very long time and very many steps to achieve what Austin has called the ‘first step’ and I will try to go through part of this process with you today.

I particularly want to illustrate how context is important in obtaining support for collections and use events that led up to the National Science Foundation Program for Advancing Digitization of Biological Collections as one specific example of what has been a process for hundreds of years. Research use is a critical component of this program and my talk should explain the “why” behind that component.

2. The themes for this talk as take away messages: information in collections is valuable for a variety of purposes and this fact needs to be emphasized, the target audience for any discussion is extremely important, framing is needed for the survival and continuing support of collections and their improvements.

3. Let’s begin with an historical perspective on support for collections

Early collections of plants happened often for beneficial or religious purposes. The ginkgo *(Ginkgo biloba)* is a living fossil – a species known from fossils and with no living relatives. They have been around since the time of the dinosaurs 200 million years ago and were close to going extinct. Very few ginkgo trees live in the wild in China, and there is reason to believe that monks might have tended these few for more than a thousand years. Ginkgo trees are often seen in temple grounds because of their Buddhist association. The trees live long and the veneration of old age was important for the monks. The seeds and leaves have medicinal properties, reputed to cure everything from asthma to hangovers. Thus a first ‘collection’ actually preserved a species from extinction because it was valued in religious use, and time and effort were spent in the preservation process. It has now become a valuable street and ornamental tree in cities and is still used for herbal medicines, although effectiveness could be questioned.

4. Moving on to early history of actual natural history collections - The world's first botanical garden and herbarium was created in Padua in 1545. In 1533 Francesco Bonafede (1474-1558) was appointed to the Chair of Lectura Simplicium at the University by the Most Serene Republic of Venice. In 1543 he petitioned for the creation of a model herbarium and botanical garden for scientific study, which was established by decree of the Most Serene Republic on 29 June 1545. Work began immediately on a plot belonging to the Benedictine Order, whose monks were probably already raising medicinal plants there. The importance of the garden and herbarium was justified because the use of plants in medicine was rather hit or miss, and the government of Venice saw the value in scientific application of cures and in the training of the physicians in the proper use and application of these medicines.

It still preserves its original layout – a circular central plot, symbolizing the world, surrounded by a ring of water. It continues to serve its original purpose as a centre for scientific research and the herbarium is still in research use.

5. Collections continued to have value for medical purposes for the next century. Hans Sloan 1660-1753–was a medical doctor, who managed to go to Jamaica as private physician to the Duke of Abermarle, but his time there only lasted 15 months. However, he discovered that a chocolate beverage was used by the locals as a medicine, but Sloan did not like it in water, so he figured out how to combine it with milk. Upon returning to England, he sold the recipe to apothecaries as a medicine. The Chelsea Physic Garden was founded by the Worshipful Society of Apothecaries of London in 1673 for its apprentices to study the medicinal qualities of plants and Sloan was a strong supporter of the Physic Garden. The milk chocolate recipe was eventually sold to the Cadbury Company which has made it much more popular than just for medical uses. Sloane used his money from being a physician to the rich and from his formula to make good investments and to support his collections. His collections were the base for the Natural History Museum in London. One can only hope the Cadbury Co. continued to support the Natural History Museum in some way, although there is no information that this happened.

6. In the 1700s, botanical collections were still used for medical purposes, but this was also a time of cooperative links between patrons and botanists, as well as some government support for discovery of resources in new lands. Linnaeus 1707-1778 had several patrons – certainly example of the system from time at university – 3 while at university, and then he became a tutor under one patron, 4 supporters on path to his doctorate, off to Clifford in Holland where several works were published as he oversaw the medicinal garden and herbarium before he landed at Uppsala as a professor. Once there he found support for many of his students to travel abroad and collect natural history specimens which were returned to the collections of Uppsala and Linnaeus for description. Many other botanists had patrons who supported their work and also universities were supporting more collections and discoveries of biodiversity.

Exploration of these new areas required cooperation among many individuals – governments and others providing exploration of the ‘new worlds’ and what they could get from them – required naturalists to bring back the treasures so cooperation was beginning to develop to conduct these explorations.

7. In the Victorian era, the study of natural science was enormously appealing to the middle classes and the public became interested in collections. At the beginning of the nineteenth century, the majority of scientists consisted of amateur scholars, who observed and recorded everyday phenomena. Societies for Natural History were formed, such as the Edinburgh Natural History Society in 1869, and these organizations provided a venue for lectures and forays into the natural environment. These hobbyists often compiled vast private collections of specimens, which they or other scientists would then catalog.

8. Also in the nineteenth century, museums, botanical gardens, and other scientific exhibitions educated and entertained the general public and introduced them to the discoveries of science. Later in the century, amateurs and trained scientists would band together to form larger scientific societies that served to provide a forum for discussion of current observations and theories. Scientists joined international natural history societies, which served as a source of funding by providing awards or medals with monetary prizes and by funding teaching positions or examination posts. In many cases, the societies replaced the private patronage of earlier periods. Science enjoyed extensive coverage in popular literature of the period. Scholars of the Victorian era have attributed this popularity to the rapid development of science and technology and the move from rural communities to cities. These changes led the public to romanticize nature and see plants and animals as exotic. Newspapers ran natural history sections, and every correspondence column became a debate over issues such as whether swallows could hibernate or whether toads could live for centuries immured in blocks of stone.

9. Jump to the end of the 20th Century and look at activities which were increasing cooperation and collaboration within and across areas.

10. Some of these received support from NSF because of the importance to advance research and they fit within the strategic plan. Promoting workshops and support for collaborations within the collections community also occurred.

11. NSF strategic plan for 2006-11 – discovery, learning, research infrastructure, and stewardship – investing in America’s future brought new possibilities for collaborative research and collaborative infrastructure projects especially as infrastructure utilized new areas for the advancement of research capability.

12. Increasingly collaborative proposals were coming in to the Biological Research Collection program and that program had not increased in budget for over 10 years, and in fact, due to a variety of cuts by Congress for the overall budget, it had lost ground. The problem, of course, was that all infrastructure programs were underfunded and overall needed new strategies to move forward.

A positive outcome from the workshops and reports mentioned earlier was a successful Research Coordination Network proposal for Collections Web with a goal to build communication and community among natural history collections. Workshops supported by this RCN have led to further emphasis on the research value of collections, tools needed for collections activities, smaller collections, education activities through collections, collections and stakeholders, and integration of collections.

Within the Federal Government, there was recognition that the U.S. collections were a heritage needing attention and that they were important for a number of reasons. This concern was expressed in the Office of Management and Budget memos to the federal agencies for consideration of priorities in their budgets.

13. The actual language from the budget memo is given here. The language continued for the budgets being developed from 2007-2009.

14. The language was an outcome of an Interagency Working Group on Scientific Collections that was formed under the Office of Science and Technology Policy in 2005. The first meeting discussed the roles of the various agencies within collection activities. NSF was represented at the meeting, but it was noted that no collections were actually held and curated by NSF. However NSF had programs that supported collections improvements and provided some support for special collections, such as polar ice cores. Thus, NSF was asked to remain as a participant in the working group to represent an interface between the federally owned collections and the nonfederal collections. When a survey of the collections was proposed to find out what the state of collections might be – NSF developed a corresponding survey for these other collections.

15. A focus group to develop this assessment questionnaire was held at the American Museum of Natural History, the questionnaire was approved by the office of Management and Budget, and the collections that had received some support from NSF over the past 20 years were sent the form. A wealth of data was received about the needs of collections, and indicated that collections were underfunded to take on many new activities.

16. The NSF survey of collections was finally fully analyzed and a summary of the findings was produced.

17. At about the same time the complete summary of findings was published and the Interagency Working Group on Scientific Collections published its report based on the federal survey of collections. All these reports had similar needs identified for collections: networking, expertise, reduce risk, accessibility improved, and new tools for exchange of data. Based on these reports, BIO asked for rationale for future requests, and several of us worked to succinctly state why we had to support these efforts.

18. The first step was to indicate where support for collections already existed in the foundation. These were spread among a number of programs, each of which made its own decisions based on the proposals presented. Thus funding decisions were independent, could be redundant, and certainly were not being analyzed for how progress overall was being made.

19. Full rationale for support of the effort was based on the two reports that had appeared from the IWGSC and the NSF survey. The main points of the rationale are summarized here.

20. Following on the input from the internal NSF group, the BIO directorate requested that its Advisory Committee discuss a variety of reports, and three of the members of the AC took on the collections reports. Their recommendations to BIO were basically three points – on slide.

21. After a series of discussions, BIO asked me to return to NSF and engage the community in next steps to accomplish the work. I did in Jan of 2010 and immediately turned to the community to ask what could be done and how it could be done. They began with a focus group, followed by blogging in the community for comments and suggestions on some points that were developed out of the focus group, and then a second focus group convened to develop a strategic plan for digitization of collections.

22. The community strategic plan was produced by the end of May 2010 and published with these levels of needs to get the job of digitizing done.

23. While the assessment was occurring, the new NSF strategic plan was approved and a major priority was to ‘transform the frontiers’ – language here. The close coupling of research infrastructure and discovery seemed to fit perfectly with the goal of forming digitization networks around major research questions.

24. In addition, the new budget authorization act for NSF – America Competes – includes language on collections accessibility for the benefit of the scientific community. And the NSF BIO budget had indicated that 10 million dollars would be devoted to starting a digitization of collections program.

25. So over two months, June and July of 2010, the NSF working group came up with a program with the main focus on collections digitization that is critical for answers to major research questions in order to fit into all the policy statements from NSF, OSTP through the IWGSC and Congress and to allow the establishment of a central organization for coordination of the projects and the data. We were up and running, and since we envisioned a 10 year program with at least 10 million per year, it will be a 100 million dollar project. And we are constantly saying this is not enough to get the job done and trying to recraft the rationale constantly to fit with new and ever changing budget priorities.

26. So let me return to the themes that I hope you remember from this example of finding support:

Information, audience, framing, and finally, based on my experience, nothing beats dumb luck

27. It is now up to the community to make this work and that means cooperation among collection personnel, researchers, and the public. We probably will not see a confluence of events this favorable again for a while.

Innovations must continue and collaborations have to increase. We need to be working beyond our comfort zones and looking to embed collections challenges in other fields of study. Cyberinfrastructure is a start, but now involving additional areas of biology and its applied fields, engineering areas, and the physical and social sciences will be needed, at least for NSF projects given the strategic plans that have appeared so far and the push from the director and the National Science Board for innovative activities at NSF. And we need to move to new ways to collaborate through virtual organizations, web 2.0, semantic web, and social networking.

The other federal agencies are working to provide accessibility and there must be integration. USGS is working on a big effort to integrate occurrence data – BISON (which has been discussed here at this meeting). And there has to be international integration of all our efforts to provide a global museum of collections across the 7 continents and the 7 seas.

So where are we going next? The challenges are great and interesting. How do we obtain data from a variety of collections, such as fish in jars and dried bones, insects in trays and genetalia on slides, plants in a variety of forms like fruits, flat specimens, wood samples, palynology slides, or birds with skins, eggs, nests, songs, and skeletons? And now that whole genome sequencing is becoming cheaper and soon will be routine, what is the appropriate way to provide the links for the molecular information that will greatly exceed our view now of simple BarCodes? How can all this disparate data be coordinated into one application searchable on many criteria? And do it on a handheld device in the field? Closer to home, how do we show measurable results in five years for politicians?

And, keeping to the theme of this BSA meeting, how do we provide for the next generation? The collections community has the challenges and now can think about assuring that the next generation truly sees this resource as critical in their lives, as other generations have, and that there are interesting questions to be solved. It is only through this outreach and education that support will continue.

28. So for the future we need to think big. How can we provide the information housed in collections as a whole, one-stop shopping, global resource? What do we want as an end result? Well, we need the collections online, we need them searchable broadly, we need to be able to annotate the information, and we want to know what others have been searching for when similar terms are used for biodiversity information. We would like to be able to conduct analyses on the results of our search for data according to set standards. We need to do more than bring dark data to light; we need to become much more inclusive and widespread. We need Collections Online with Scientific Meaning of Organisms for Society.