

iDigBio Response to the 2017 Report of the External Advisory Board

March 29, 2018

Summary

iDigBio's Project Director, Project Manager, Principal Investigators, and Project Evaluator met in-person with the External Advisory Board (EAB) on November 1, 2017, at the ADBC Summit held in Gainesville, Florida. The meeting was observed by Roland Roberts, the NSF Program Director for iDigBio. As a result of the meeting (and as required by iDigBio's Cooperative Agreement with NSF), the EAB provided a report documenting recommendations to iDigBio regarding sustainability, with consideration given to operating after the NSF-ADBC program, and, more specifically, increasing data use by research as well as education-outreach activities.

This document consolidates the recommendations in the 2017 Report of the iDigBio External Advisory Board and provides iDigBio's responses to those recommendations. To facilitate internal discussion and action, issue tickets have been created iDigBio's Redmine project management system. A parent issue ticket (#2541) was created for the report itself to which subtasks were assigned for each recommendation in the report:

- Issue #2542: Specific plans/strategy to complete current goals and ensure sustainability
- Issue #2543: Increase engagement with the TCNs to find ways to expand data use
- Issue #2544: Work together with TCNs to seek NSF support of Postdoctoral research using digitized data
- Issue #2545: Enumerate institutional affiliations of data users of portal
- Issue #2546: Quantify portal data use across research fields
- Issue #2547: Gather a complete picture of ADBC data use
- Issue #2548: Gather information about institutions, collections, and specimens remaining to be digitized
- Issue #2549: Establish targets for data mobilization and ingestion
- Issue #2550: Identify cause of decline in rate of data ingestion
- Issue #2551: Assist with sustainability of the national digitization effort
- Issue #2552: Develop a detailed sustainability transition/business plan
- Issue #2553: Secure ongoing core support from a government agency
- Issue #2554: Develop specific goals and priorities for the remainder of ADBC funding
- Issue #2555: Develop closer ties with the international community

Response

iDigBio appreciates the time, effort, and concern of the EAB in ensuring the success of iDigBio and the national digitization effort. iDigBio found the 2017 report to be quite helpful, particularly in focusing attention on priorities for the remaining years of ADBC funding. iDigBio will continue to partner with the EAB on the recommendations voiced in the report.

Looking toward a sustainable model for maintaining the digitization and data mobilization successes of the past seven years, the iDigBio PIs have established a “Committee of Five” to develop draft strategic and implementation plans to guide the project towards and into post-ADBC funding. These plans will outline iDigBio’s transition into a U.S. GBIF node, the transition of selected functions (Research and E&O) to the Florida Museum, and details for continuing iDigBio’s project management, data mobilization, portal, and training and workforce development functions. The committee is charged with completing its work by Summit 2018, with periodic reports to the Executive Committee. Members of the committee, representing all domains of iDigBio, include: Gil Nelson (Chair), Shari Ellis, Renato Figueiredo, Joanna McCaffrey, and Molly Phillips.

iDigBio has also made good progress in Research Use, Sustainability, and Education & Outreach, covering many aspects of the recommendation in the report:

RESEARCH USE OF DIGITIZED DATA

iDigBio has continued to work with the collections community to digitize and mobilize specimen-based data as quickly and effectively as possible. Nearly 110 million specimen records (representing ca. 330 million specimens) with 24 million associated media records are now searchable through iDigBio. There has been an increasing emphasis on improving data quality, access, and use in research through a number of initiatives, including development of easier access to research tools and software, developing collaborations, and hosting or otherwise attending training workshops, webinars, symposia, and other events covering many digitization and data-related topics to inform researchers. iDigBio PIs and staff gave multiple presentations throughout the year and across the U.S. to promote digitization and data use. In addition, iDigBio has catalyzed an annual conference highlighting digital data in biodiversity research (U. Michigan in 2017, Berkeley in 2018, and Yale in 2019). iDigBio is working with GBIF and the EAB to improve data use metrics and to ensure appropriate attribution of data use.

Research workflows are being developed to illustrate how data can be used in new research applications. User-friendly interfaces are being developed to expose new capabilities of iDigBio APIs through new visualizations in the specimen portal, new appliances, or integration with third-party applications software. A major initiative to improve access to research tools involved collaboration between iDigBio and Encyclopedia of Life (EOL) and led to the development of two tools using iDigBio biodiversity data and other data sources: (1) Effechecka, a taxonomic checklist generator which uses spatial and trait data, and (2) FreshData, which enables researchers to receive notifications when new biodiversity data, based on a particular query string, are available from various data sources, such as iDigBio. Code for both these applications are available via GitHub. The collective effort has been termed GUODA – Global Unified Open Data Architecture (<http://guoda.bio/>). The goals of the continued

collaboration and development are to share data use cases, best practices, infrastructure, code, and ideas that can be done by analyzing large open-access biodiversity datasets.

Another important initiative to increase specimen data use is to link genetic data, widely used in phylogenetics, to specimens in collections, which will greatly increase verifiability of published genomic and other genetic data as well as use of specimen data. iDigBio is working on this with GenBank, BOLD, TDWG, several TCNs and others interested in tracking genomic data. Likewise, iDigBio is working with the Spatial Information Systems Lab at SDSC, to develop an extension for exploring large databases, support domain standards, interface with taxonomies, and make it available through an analysis platform.

Lastly, iDigBio is working to expand its user base through a number of initiatives to increase use of collections data by conservationists and ecologists, which has included meetings with NEON, a symposium at ADBC Summit VII, and meetings with attendees at ESA and SPNHC.

SUSTAINABILITY

iDigBio's transition into a sustainable program beyond the 10 years of NSF funding will depend on (1) its importance to the collections and research communities, (2) the components to be sustained, and (3) a combination of institutional and external funding sources. Sustainability of iDigBio and the national digitization effort depend heavily on large-scale participation by the collections and research communities. The number of collection-holding institutions iDigBio is working with has now increased to 708, making iDigBio, along with SPNHC and NSCA, a leading advocate of the biodiversity collections community. iDigBio's position has been further strengthened by the continual workshops, webinars, symposia, and other events on digitization and data-related topics hosted and facilitated by iDigBio.

Increased support at UF is progressing along several fronts, including discussions with the UF Libraries Data Management Services to deliver a consistent message across the institution about the need for research data management. Larry Page has been participating in discussions with representatives from several key systems, including iDigBio, EOL, GBIF, CoL, and ALA, on how to coordinate and link efforts globally by developing a collaborative e-infrastructure to share data, increase efficiencies between existing platforms, and enable better-informed research and environmental decision-making, which helps promote sustainability on a broader level.

In 2017, iDigBio signed a Memorandum of Understanding (MoU) to become an Associate Participant in the Global Biodiversity Information Facility (GBIF), increasing its visibility and prominence in the international community. iDigBio named Greg Riccardi as the Head of Delegation for the iDigBio GBIF node (<https://www.gbif.org/participant/375>) with David Jennings as the Node Manager and Joanna McCaffrey as the Deputy Node Manager. In addition, David Jennings was elected as the regional representative for all North American GBIF nodes (<https://www.gbif.org/the-gbif-network/north-america>). As such, he will participate in the Nodes Steering Committee, which shares information about the status and best practices of Participants' nodes and serves as an advisory committee that makes recommendations to the Governing Board, the Executive Committee, the Science Committee, and the Secretariat concerning relevant issues for the nodes.

EDUCATION AND OUTREACH

E&O efforts in the past year focused on expansion and coordination of K-20 formal education and informal science activities across TCNs and others in the collections community through programs that emphasized digitization, biodiversity informatics, and the value of collections data for understanding biodiversity. Training of students built upon successful elements in Phase 1, including cyber-enabled courses that connect students across institutions. Lifelong learners are being targeted with the goal of promoting participation directly in the national digitization effort through close integration of Citizen Science with digitization activities. iDigBio is continuing its efforts to broaden participation in biology by sponsoring career workshops targeting underrepresented groups.

iDigBio personnel participated in more than 20 symposia, workshops, and short courses, including six workshops or short courses focused on training or creating a product (e.g., developing best practices) and at least fifteen events associated with professional meetings. In addition, iDigBio did outreach at three conferences using the booth, Libraries of Life cards, bookmarks, and other marketing materials. Moreover, iDigBio engaged in two events that targeted K-12 teachers and students.

iDigBio held more than 30 webinars, one of which was organized by the Education and Outreach working group, four were coordinated by the partner FOSSIL Project, five were coordinated by the Biodiversity Heritage Library (BHL), and the remaining webinars covered a wide variety of topics including Darwin Core, Symbiota, Arctos, and georeferencing.

iDigBio engaged in the global WeDigBio event held in October both as a partner and as a participant with events at FLMNH and in Tallahassee. iDigBio is partnering with a number of organizations including Aim-Up!, BLUE, QUBES, BCoN, and KURATOR on a Faculty Mentoring Network developing modules for undergraduate biology education.

Lastly, iDigBio currently has 1,154 newsletter subscribers, 2,222 followers on Facebook, and 2,704 Twitter followers.