2019 Report of the iDigBio External Advisory Board

November 21, 2019

iDigBio External Advisory Board (EAB) membership:

- Neil Cobb, Northern Arizona University (Chair)
- Linda S. Ford, Harvard University
- Donald Hobern, Global Biodiversity Information Facility
- Jason Knouft, Saint Louis University
- Barbara Thiers, New York Botanical Garden

This report is based on the following:

- In-person meeting with all EAB members, iDigBio key personnel, along with Roland Roberts (NSF observer) at ADBC Summit IX in Gainesville, Florida.
- Presentation by Gil Nelson at EAB in-person meeting
- Presentation prepared by David Jennings documenting iDigBio's progress in FY8

The EAB is charged with providing recommendations to iDigBio regarding sustainability, with consideration given to operating after the end of the NSF-ADBC program, and more specifically increasing data use by research as well as education-outreach activities. In 2019, the EAB worked as a single committee to assess sustainability and core functions of iDigBio, including data use.

The EAB reviewed previous documents, the 2018 EAB report and meeting materials, and the information provided in association with the 2019 in-person EAB meeting. For 2019, the EAB does not have new specific recommendations and has chosen to update and reiterate its recommendations from the 2018 report.

General EAB Recommendations

Below are the key points based on the 2018 & 2019 meetings at the ADBC Summit.

1) The EAB fully supports the creation of a reconstituted EAB that can provide guidance to iDigBio in a post-ADBC period. The new EAB may include members of the existing EAB membership, but iDigBio will also solicit complementary expertise to address new challenges.

2) The EAB continues to encourage iDigBio to emphasize further refinement and prioritization of the plans outlined in the Committee of Five (CoF) document and the 2019 presentation.

3) The EAB suggests that iDigBio continue to frame the action items listed in the CoF report in the context of the (seven) "critical functions" or any refined version of these.

4) We encourage iDigBio to continue to refine the "critical functions" as key services offered to the community or as user stories rather than as broad areas of function. Furthermore, iDigBio should present a business plan and sustainability strategy for each such service, recognizing that the strategy may be for collections and/or other parties to assume responsibility for some of these services in the longer term. iDigBio should identify an achievable number of core functions to be sustained by iDigBio itself.

5) International partnering leading to a globally connected pipeline will continue to be the most significant and challenging aspect of iDigBio planning.

6) We are confident that iDigBio is well positioned to complete its ADBC directive.

In summary, the EAB supports the direction and planning activities initiated by iDigBio with the ADBC Strategic Implementation Plan. We strongly encourage the iDigBio to continue these efforts with more specific plans for how the goals of an ADBC-like effort will be sustained after 2021, with careful consideration for prioritizing functions and collaborations both with the ADBC community and beyond to create a global biodiversity pipeline.

Specific Recommendations on CoF Planning Document & 7 Core Functions

The Committee of Five Document and planning Stages

The CoF document is titled "ADBC Strategic Implementation Plan", but the document appears to focus on iDigBio activities. If this plan encompasses all of ADBC, rather than just iDigBio, it would be helpful to see more detail on how iDigBio would engage the collections' institutions in developing and "owning" the long-term, sustainability strategy. There are currently items listed that relate to supporting the data systems (e.g., Symbiota, SCAN, Specify, Arctos) and existing initiatives (e.g., data carpentry, Darwin Core Hour), but much less specificity around institutional engagement.

The structure of the CoF document is focused on processes and mechanisms rather than goals and deliverables. We suggest that the ADBC Strategic Implementation Plan should instead be structured around a clear view of a set of services that iDigBio should assure for the long-term benefit and support of the collections community.

The Seven Critical (i.e., Core) Functions identified in the presentation delivered by Gil Nelson at the EAB meeting in 2018 (listed below) offer an effective framework for presenting these services. However, the EAB believes that the critical functions as currently listed are too generic, could probably be restructured as a smaller set, and should be more tightly defined as a set of key services that iDigBio is offering to the community. It may be most effective to present these services as a set of user stories – e.g., "A collection manager uses service A from iDigBio to achieve X, Y and Z". Such a view of the iDigBio offerings would give clarity and confidence for all stakeholders about the current scope and vision.

Critical Functions and Planning Stages

iDigBio Core Functions
1. Workforce training/community building
2. Education, outreach, and diversity (EOD)
3. Data mobilization
4. Data support/troubleshooting
5. Special access to specimen data
6. Encouraging/tracking research use of data
7. Continuous program evaluation/documentation

iDigBio Planning Stages

Stage 1: 2019 to 2021 (August)
Stage 2: 2021 to 2024 (post ADBC funding)
Stage 3: Post 2024

Defining this set of key services, which should focus on the capabilities, both technical and social, that are being supported by iDigBio during the current program, will allow the ADBC Strategic Implementation Plan to offer a future sustainability vision for each of these services. In some cases, the future plan may be to secure adequate resources within iDigBio to maintain the service at the level needed by the community. In other cases (e.g., some aspects of workforce development), the goal may be to ensure that the collections community itself has the capability to maintain the service following the end of the current funding period. In yet other cases, the sustainability plan may reduce long-term resource needs through planned closer collaboration with GBIF, ALA, or other external parties to ensure cooperative provision of a shared service.

An ADBC Strategic Implementation Plan of this nature could support fruitful discussion with collections, with other research infrastructures, with NSF and with other funding sources to secure an optimal approach to sustainable digital management of US collections. This approach can also clarify iDigBio's unique niche as the catalyst for coordinating efforts and resources to ensure that researchers have the capability they require.

The Plan can also present, for each service, a timeline for how each service is to be maintained and improved over Stage 1 and Stage 2. By the conclusion of Stage 2, iDigBio and the ADBC community should be in a maximally sustainable form for Stage 3. Regarding Stage 3, iDigBio should provide a general perspective for the period around 2024-2028, regarding what future investments might possibly build on the foundations of the ADBC program and how these might either depend on the sustainability of the critical iDigBio services or else host components of the next stage of a persistent iDigBio capability.

The 34 action items listed in the CoF document should be linked to the critical functions. They could be further developed in more detailed plans defined by the seven critical functions. For each of the services identified as necessary on behalf of the collections community and researchers through 2024 and onwards, the focus should be answering several important questions, including:
1. What is the plan, and what resources are required, for maintaining and improving this service during Stage 1?
2. What can be done to maximize the sustainability of the service by the end of Stage 2?
3. What are the options for provision of this service in Stage 3? This raises a suite of questions the EAB feels are critical for long-term sustainability of the ADBC effort.
   A. Does this service need to remain part of a continuing iDigBio/post-iDigBio centralized infrastructure?
   B. Could it be maintained sustainably as part of the operations of collections that have benefited from ADBC?

C. Is it possible that partners such as GBIF can provide redundancy to guarantee persistence, or some or all services be transferred to an interested institution on an interim or permanent basis?

The Internationalization section is about a strategic approach to collaborate on developing a global pipeline. This would involve collaborating with GBIF, ALA and other programs and allows iDigBio to leverage existing efforts more effectively. This has been the area where iDigBio has made the most progress in terms of developing a strategic plan for specific collaborations with GBIF, ALA and other programs. Developing closer ties with the international community is likely to reduce redundancy allowing more focus on other areas that will lead to a stronger cyberinfrastructure better able to serve researchers the data they need. The contents of the Internationalization section then become an explanation of plans during Stages 1 and 2 to evolve some of the existing iDigBio services to a more persistent and sustainable shared model. The reason and justification for developing partnerships is to enhance iDigBio infrastructure while also improving its sustainability. Partnering internationally is the proposed mechanism to do this and the "Internationalization" section is good but should be re-named to better fit iDigBio and ADBC goals.

The remaining two sections of the CoF document appear to be more detailed action items that encompass the core functions. The second section in the CoF (Education, Outreach, and Diversity (EOD)) is an expansion of the first and second "Core Functions", and section III (Preservation) is about sustaining Core (i.e. Critical) Functions.

Seven Critical Functions

The EAB members continue to support iDigBio exploring ways to integrate the seven core functions. Specific possibilities included combining #3 (Data Mobilization) and #4 (Data support/troubleshooting), and both of these fit well with #6 (Encouraging/tracking research use of data). The core functions #1 and #2 could be combined into a 'Training, education, and outreach focus'.

Critical function #1 (Workforce training/community building), #3 (Data Mobilization) and #4 (Data support/troubleshooting) are the most fundamental of functions; furthermore items 3 & 4 can be viewed as different aspects of the same function. Additionally, without number 6, the data mobilization effort doesn't have any impact and/or its impact can't be demonstrated, thus it might be appropriate to integrate core function # 6 (Encouraging/tracking research use of data). Funding for iDigBio, partnerships (with TCNs, scientific societies, GBIF/ALA, educational initiatives, federal agencies and others), EOD, and new funding proposals all become activities that can be prioritized and pursued based on their likely contributions to sustaining, growing and evolving the essential functions. We do not feel that core function #7 is necessary to delineate as a separate function, but could be acknowledged in the other core functions. The EAB is not

requesting the elimination of any core function; it would likely be premature to do so without additional rounds of planning.

There is some concern that even the "Seven Critical Functions" are too general. These are areas of activity rather than defined services or capabilities that iDigBio offers to the community. It could be much more powerful to define the key services that iDigBio offers to its stakeholder community, perhaps defined as, or at least clarified by, user stories (as noted earlier). The list of core functions should provide a sharper, crisper set of definitions of what iDigBio delivers for each of these separately. For example, what are the critical functions of iDigBio concerning Data mobilization (#3)? Which of the following are included in iDigBio's critical function: documentation, hands-on support, helpdesk, training courses, publishing tools, and/or data hosting services? Is #4 (Data support/troubleshooting) actually a separate function from #3 (Data mobilization)? For some of these, the need to clarify is even greater - especially EOD.

Once there is a clear view of the actual services that iDigBio could maintain going forward, it should be possible to rank these into Critical, Desirable, and Optional. This type of ranking would offer more flexibility as funding potentially varies, and it demonstrates a priority analysis of the functions that could be provided to the collections community. At this stage, it may be wise to identify the most essential core of iDigBio rather than committing to do everything indefinitely.

iDigBio
Integrated Digitized Biocollections

Minutes of External Advisory Board Meeting on October 1, 2019

Date/Time: Tuesday, 1 October 2019, 1:30-3:00 PM Eastern

Attendees: • **EAB:** Neil Cobb; Linda Ford; Jason Knouft (via Zoom); Barbara Theirs
• **iDigBio:** Renato Figuerdio; Jose Fortes; David Jennings; Bruce MacFadden; Austin Mast; Gil Nelson; Greg Riccardi; Pam Soltis; David Blackburn
• **NSF Observer:** Roland Roberts

Discussion Items:

EAB Restructuring

- Global representation on the board is important. Next few years is really critical in keeping up with international developments. Need Global perspective. Need to be working in parallel and not without communication.
- Many folks on the specimen network would be great to have involved in the board.
- Have someone on the board that understands the board history.
- We need new stakeholders that are outside our normal stakeholder group. Someone who represents industry, public health, etc. All of these people can use our data. It is important to have a broader perspective – with a different audience.
- Need someone who represents the INCLUDES network at NSF on the board. We should integrate more fully with INCLUDES.
- Advisory board members should be given action items throughout the year. This would serve as a way to get more out of them during the year as opposed to just once a year.

SABI Proposal

- TCNs and PENs will be funded up to year 10. Some form of iDigBio is needed to provide services to that community. The goal is to preserve services for new TCNs, existing TCNs with termination dates past 2021, CSBR projects that include digitization, and non-NSF-funded institutions that are digitizing and mobilizing data.
- iDigBio is in the process of writing a SABI proposal to help sustain the services of iDigBio.
- The Sustained Availability of Biological Infrastructure program (SABI) supports the continued operation of extant infrastructure that will advance basic biological research.
- iDigBio is planning to submit the proposal in the first quarter of 2020.

Internal Advisory Board

- The concept of an Internal Advisory board needs to be refocused as well. It is not effective in its current state.

- There are a lot of people who do not attend the meetings. They do not provide input. They don't feel a part of the community. If we are moving towards globalization, we can't leave the TCN community out. The IAC meeting needs to be restructured and made more important to the TCN community.
- IAC Meetings should be like the Summit online…but isn't! There have been ways in the past to increase attendance that haven't worked.
- Maybe a solution could be to replace the IAC with working groups or webinars that can engage them more.
- Maybe there should be something more to the meetings that make them beneficial to the attendees…instead of them just reporting.
- It is important for the new TCNs to help them to be more open across the ADBC.

Announcements and News

- Jillian Goodwin has a new position as Conference Manager at UF. She coordinates iDigBio's annual conferences including the Digital Data conference and ADBC Summit. She also works with the ADBC community to develop workshops.
- Libby Elwood is the Global Communications Manager. The global communications manager will provide leadership in organizing, providing structure and maintaining momentum for our globalization efforts.
- Erica Krimmel was hired at FSU as Digitization Resource Coordinator. She is responsible for organizing and leading working groups and workshops, improving resources for digitization training and workforce development, and fostering collaboration.
- ACIS has almost completely re-staffed. New guys are awesome and have integrated nicely. The staff has made important improvements in portal response time and capacity.
- Read more: Response to 2018 EAB Recommendations

iDigBio Stats and Charts

- iDigBio Progress in FY8

Other news

- Bruce will step down as PI. Looking to recruit someone who has focus on Education and Outreach. It likely will be David Blackburn.
- Shari Ellis has retired. iDigBio will most likely hire an in-house staff member to take over her duties