Minutes of Internal Advisory Committee Meeting on February 7, 2018

Date/Time: Wednesday, February 7, 2018 - 2:00 PM Eastern

TCN Attendees:

SoRo	Ryan Allen	InvertEBase	Petra Sierwald
PALEONICHES	Bruce Lieberman	GLI	Ken Cameron
EPICC	Erica Clites	SERNEC	Michael Denslow
LepNet, SCAN	Neil Cobb	MiCC	Andy Miller
Cretaceous World	Bruce Lieberman	MAM	Cindy Skema, Anne Barber
oVert	David Blackburn, Oona Takano		

iDigBio Attendees:

Administration	Cathy Bester	Digitization & Training	Deb Paul, Gil Nelson
Education & Outreach	Molly Phillips	Cyberinfrastructure	Joanna McCaffrey
Research Use		Assessment & Evaluation	Shari Ellis

NSF Attendee: Reed Beaman

Facilitator: Neil Cobb

Meeting Location: Building 105 Room 310 and http://idigbio.adobeconnect.com/iac/

Recording: http://idigbio.adobeconnect.com/p6ulx71pdhq3/

Opening Remarks

The frequency of the IAC meetings has been reduced from 6 to 4 times per year, as these meetings have not been very productive on either side (iDigBio or TCN). These meetings need to be more relevant, with TCN reps playing a larger role. iDigBio staff may need to stand away and let the TCN representatives determine how these meetings are structured. In Neil's view, the purpose is to structure increased TCN involvement with more substantive items discussed.

Standing Items

Molly is looking for research topics to include in the Spotlight Newsletter, maybe write a blog on something interesting. Have you subscribed yet? https://www.idigbio.org/newsletter-subscribe

Agenda Items
Google Analytics

A more complete assessment of activity supported by the ADBC program is needed to show success and for possible future funding and sustainability opportunities. (iDigBio has positioned itself well to be a GBIF node

which will provide some basic support for sustainability, however at this time there doesn't appear to be any TCNs who are planning for post-ADBC funding.) The entire community needs to pool usage statistics for web sites and portals using Google Analytics when possible. Google analytics data is a simple way to track a lot of activity supported by the ADBC program. Neil logged into Google Analytics Data Studio and ran across recordsets that Alex had make for iDigBio which were quite useful, allowing him to compare SCAN with iDigBio over the past year for long term trends. Alex set up the reporting structure which took Neil's data and put it into the same format as iDigBio's. iDigBio has separate stats for the web site and for the portal; the SCAN portal is getting more use than the iDigBio portal currently. More effort needs to be made in advertising the iDigBio Portal to increase usage. Neil reported that the Data Studio part of Google Analytics was a wonderful find for him, providing valuable information. Andy requested that someone write up a one page report on how to get data report highlights from Google Analytics which would provide assistance for the community for data gathering. Neil doesn't feel that he is knowledgeable enough to write this however perhaps Alex could assist or perhaps someone with another TCN. Neil has reached out to Eva Huala (EAB member), she may be of some assistance in this area.

The link to the Google Analytics Data Studio:
https://datastudio.google.com/u/0/navigation/reporting

Neil would like to know: who has project web sites/portals, are you collecting Google Analytic stats, and if not, are you willing to collect and share Google Analytic stats? Email neil.cobb@nau.edu If everyone provides statistics for their respective project web sites and portals that we can all share, NSF would be pleased. Neil suggested that up to five metrics are reported quarterly as part of the existing reports that are submitted to iDigBio; metrics will be reported over the last year (or as long as the sites have been gathering Google Analytic data) as monthly totals/means and a one-time summary of Google Analytics metrics gathered since it was implemented on your site. (Michael did note that he has attempted to embed Google Analytics code into his sites with no success and no support from his institution.)

Data Quality – Geography

Joanna discussed a few key areas of data that if implemented by TCNs will increase the quality of data being served.

Flag reference: A complete guide to the flags is found here: https://github.com/iDigBio/idigbio-search-api/wiki/Data-Quality-Flags

To view the flags on your dataset, begin here with the publishers list:
https://www.idigbio.org/portal/publishers

General data practice: Put all data that might be in doubt or of the 'verbatim' nature into verbatimLocality field. Use the continent, state/province, county, countryCode for values that might be in an ISO authority list or from reserved vocabulary. Avoid data tics (e.g., *, ?, parentheticals, place holders) and modifiers (e.g., for Putman Co. in the county field, just put 'Putnam'). Another example is taxonRank. If using 'sp.' or 'sp. nov.' in specificEpithet field, the taxonRank value is 'genus', not 'species'.

Your dataset: Find your dataset in the list and then look at the 'Data Corrected' tab. If you click on any flag, you will get a list of the records that raised that flag. Look at anything 'geo'-related for what should be easy repairs back at your data source.

Repair and re-publish: Remember to re-publish your data and we'll get the updates on our next data update.

Role of TCNs in defining post-ADBC funding

Does anyone want to participate in developing a strategy for future funding and if so is there critical mass to create an ad hoc TCN group to represent TCNs?

The data provider community can play a significant role, is anyone interesting in forming this ad hoc TCN group to help define something comparable to the NIBA document and to seek funding after 2021? Petra and Bruce are interesting in helping; Anne and Ryan may be as well depending on the time commitment. Ken has indicated that once his TCN is over, he will be finished due to his TCN theme that lacks geography and taxon focus. In the end his TCN didn't network so much as a scientific community and as such he doesn't see this moving forward. In regards to BCON, Gil mentioned that during the data conference in Berkeley (June 2018), BCON will organize a discussion session on what's next for ADBC funding, digital data, and data use. Unsure regarding virtual participation however the session will be recorded for those who are unable to attend the conference. If you are able to attend, feel free to join the discussion as input from different perspectives is welcome.

How should TCN meetings be structured?

Given that iDigBio TCN meetings are now quarterly, how do we structure the meetings? TCN feedback would be very useful. Should iDigBio PIs participate in the TCN meetings?

It was suggested that the meeting name should be changed from IAC to something more appropriate. Neil would like the TCNs to think more on how to best structure these meetings due to complaints as there is nothing to prevent TCNs to make the meetings more useful.

Suggested topics for May 2018 meeting

TCNs should structure and facilitate these meetings – Neil facilitated this meeting, another TCN should lead the next meeting in May. David is the chair of the meeting to bring more of a project management focus, and as such should be included in this discussion. Joanna requested that a data quality corner be included at each future meeting. Erica would like a data portal discussion, in particular when will the next version of the portal be up and ready? Bruce agrees it's a good idea to include data quality corner, and to hand the meeting off to the TCNs for each meeting. Neil created the agenda based on what he was interested in discussing for this meeting. Bruce is willing to lead the May IAC Meeting on May 2, 2018 and will send out an email for agenda items.

Discussion Items

Deb Paul announced the GBIF funding opportunity: GBIF call for nominations for young researchers award – nomination-based, graduate or undergrad student using collections data/gbif data, which will advertise the importance and research use of collections. For more information, visit: https://www.idigbio.org/content/nominations-gbif-young-researchers-award-2018

Neil also mentioned that Mary Barkworth is finalizing a proposal to GBIF to get Pakistan plant data, opening it up to try to bring in people from Pakistan. The community should be targeting parts of the world that we want that ADBC to reach out to and encourage collaborations. GBIF is a nice way to get seed money to foster these activities.

Ryan informed the group that his (SoRo) TCN is struggling with locating a new source for lightboxes (and bulbs) with OR Tech/MK Digital direct out of business. Cindy has helpful information if anyone is interested, regarding a contractor who builds lightboxes. Anne, Cindy, and Ryan will continue this discussion via email.

Deb said that as a result of GBIF efforts in Russia, there is interest for Russian language documents and materials to be placed on the iDigBio web site which will assist their digitization efforts.

Meeting Adjourned: 2:55 pm